DOCUMENT RESUME

ED 399 578 CS 509 319

AUTHOR Morreale, Sherwyn P.; And Others
TITLE Large Scale Assessment of Oral Communication: K-12 and Higher Education. Second Edition.
INSTITUTION ERIC Clearinghouse on Reading, English, and Communication, Bloomington, IN.; Speech Communication Association, Annandale, Va.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO ISBN-0-944811-19-1
PUB DATE 96
NOTE 106p.
AVAILABLE FROM Speech Communication Association, 5105 Backlick Road, Building F, Annandale, VA 22003.
PUB TYPE Book/Product Reviews (072)

EDRS PRICE MF01/PC05 Plus Postage.
DESCRIPTORS Elementary Secondary Education; Higher Education; *Oral Language; *Speech Communication; *Test Content; *Test Reliability; Test Reviews; *Tests; Test Selection
IDENTIFIERS Alternative Assessment; *Large Scale Assessment

ABSTRACT

This publication identifies and examines existing assessment instruments in oral communication for K-12 and higher education, and abstracts and describes assessment instruments and systematically reports their availability and interest to scholars, teachers, and administrators. It also provides background research needed to develop oral communication assessment instruments designed to encourage the development of new instruments. The document begins with an editor's foreword and a preface that relates alternative assessment, test fairness, and assessment utility to communication. It then presents an introduction to assessing oral communication, discusses topics including why oral communication skills should be taught, oral assessment for classroom decision-making, issues in assessing oral communication, locally developed instruments, and research and development priorities. The publication then discusses trends in K-12 state standards and assessment efforts, performance assessment, Speech Communication Association standards for assessment, and identification of measures. It then presents reviews of 16 oral communication assessment instruments for grades K-12. Next, it addresses the process by which instruments for higher education were identified and selected for inclusion, as well providing ideas for using the instruments. The final section presents reviews of 45 higher education instruments, divided into sections containing instruments intended for skills and behavioral assessment, or instructional and informative measures. (RS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 399 578

Large Scale Assessment of Oral Communication: K-12 and Higher Education

Second Edition

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Sherwyn P. Morreale, *University of Colorado, Colorado Springs*
Philip M. Backlund, *Central Washington University*

in collaboration with

Judith M. Dallinger, *Western Illinois University*
Ray J. Fenton, *Anchorage School District*
Ellen A. Hay, *Augustana College*
Neil J. O'Leary, *Dos Gatos Enterprises*
Don L. Rubin, *University of Georgia*

CS509319

Speech Communication Association
5105 Backlick Road, Building F
Annandale VA 22003

ERIC Clearinghouse on Reading, English, and Communication
Bloomington IN 47408

BEST COPY AVAILABLE

Acknowledgments

We are indebted to the editors and reviewers of the volume's first edition. Many of the ideas, and in fact some of the language, used in this second edition are borrowed from the first. Most particularly, we acknowledge the editors of the first edition, Don Rubin and Nancy Mead, who worked in collaboration with John Daly, Patrick Dickson, James McCroskey, and Janis Patterson. We also would like to acknowledge the support of Roy Berko in the Speech Communication Association (SCA) National Office; Gus Friedrich, who served as SCA Publication Series Editor; and David Huggins, who researched references and edited this volume for us in its final stages of review for publication. Finally, we are indebted to the external reviewer of this second edition, whose support and comments made this volume possible.

Published by the Speech Communication Association and the ERIC Clearinghouse on Reading, English, and Communication

Production Editor: Ted Spencer
Cover Design: Shawn Richardson

Copyright 1996 Speech Communication Association. All rights reserved.

Brief portions of material in this publication may be copied and quoted without further permission with the understanding that appropriate citation of the source of the excerpt will be included in such copying.

A limited number of copies of brief portions of material in this publication may be made for scholarly or classroom use if: 1) the material is distributed without charge or no fees above the actual duplicating costs are charged; 2) the materials are reproductions, photocopies or copies made by similar processes and not reprints or republications; 3) the copies are used within a reasonable time after reproduction; 4) the material includes the full bibliographic citation; and 5) the following statement be clearly displayed on all copies: "Copyright by the Speech Communication Association. Reproduced by permission of the publisher."

A copy of this statement serves as the Speech Communication Association's official permission for using material for scholarly or educational purposes under the above conditions.

This permission does not extend to situations in which 1) extensive amounts of material are reproduced or stored in an electronic or similar data retrieval system 2) a fee above actual duplicating costs is charged or if there exists a reasonable expectation of profit, or 3) the material is reproduced or reprinted for other than scholarly or educational purposes. In such cases, permission must be obtained prior to reproduction. A reasonable fee will generally be assessed for commercial copying or reproduction.

Requests for permission to reproduce should be addressed to Publications Manager.

Speech Communication Association
5105 Backlick Rd Bldg #F
Annandale VA 22003.

Printed in the United States of America
96 97 98 99 00 5 4 3 2 1
ISBN 0-944811-19-1

Table of Contents

Acknowledgments ii

Editors' Foreword v

A Preface Relating Alternative Assessment, Test Fairness, and Assessment Utility to Communication 1

Alternative Assessment and Communication 2
Test Fairness and Communication 3
Assessment Utility and Communication 3
References 4

An Introduction to Assessing Oral Communication 5

Why Teach Oral Communication Skills? 5
Oral Assessment for Classroom Decision-Making 6
Oral Assessment for Programmatic Decision-Making 6
Oral Assessment for Certifying Student Competencies 6
Background to This Publication 7
Objectives of This Publication 8
Procedures for Reviewing Assessment Instruments 8
Issues in Assessing Oral Communication 10
Content of Assessment Instruments for Speaking and Listening 10
Criteria of Assessment Instruments 11
Responses and Scoring Procedures Used in Assessment Instruments 12
Administrative Feasibility of Assessment Instruments 12
Target Populations and Potential Sources of Bias 13
Locally Developed Assessment Instruments 13
Speaking 14
Listening 15
What Should Be Tested? Research and Development Priorities 15
Are the measures you choose to use or to develop valid? 16
Are assessment instruments reliable? 16
What Measurement Techniques Are Presently Available? 17
Are Assessment Instruments Inherently Susceptible to Group and other Biases? 17
Should Communication Competence Be Assessed?—A Postscript 18
Conclusion 18
References 19

Oral Communication Assessment Instruments for K-12 21

Trends: K-12 State Standards and Assessment Efforts 21
Table: Status, by State, of Speech Communication Assessment Programs 22
Performance Assessment: A Growing Emphasis at K-12 22

SCA Standards for Assessment 22
General standards 22
Assessment content 23
Assessment instruments 23
Procedures and administration 23
Identifying Measures 23
Selecting Measures 24
Psychometric Properties 25
Conclusion 25
References 25
Selected Bibliography 26
Reviewers' Note 26
Editors' Note 26
Table: Oral Communication Assessment Instruments: K-12 27

Assessment Instruments for K-12 30

Oral Communication Assessment Instruments For Higher Education 45

Identifying and Selecting Measures 45
Using Measures 46
References 47
Editors' Note 47
Table: Skill and Behavioral Assessment Instruments 48
Table: Instructional and Informative Assessment Instruments 49

Assessment Instruments for Higher Education 51

Editors' Foreword

Sherwyn P. Morreale and Philip M. Backlund

The challenge of editing this second edition of *Large Scale Assessment* was an informative and weighty experience. We learned a great deal, and we have come to recognize the strong need in our discipline for this revised volume. Editing this volume involved making several key decisions regarding coverage and structure. In order to help the reader understand why certain instruments are included, we describe our thinking on such decisions.

Our first guiding principle was to try to be inclusive rather than exclusive, particularly with regard to instruments for 2- and 4-year colleges and universities. Perceiving a need in the communication discipline for more information on more assessment tools, we attempted to *include* tools rather than *exclude* them. In doing so, note that the reviewers of instruments for Part II and Part III varyingly established and respected certain standards for inclusion of any tool. Those varying standards are articulated in the introduction sections to the compendiums in Part II for K-12 and in Part III for 2- and 4-year schools. For example, the adherence to standards for inclusion of instruments for K-12 tended to exclude self-report measures as well as tools that assess predisposition, or communication traits or styles. By contrast, the adherence to standards for inclusion of instruments for 2- and 4-year schools tended to include such measures.

Complete consensus was difficult to obtain regarding which instruments to include and what types of information to provide in each review. Inclusion of an instrument in this volume does not represent an endorsement of it by the editors, the reviewers, or by SCA. Our goal is to present sufficient information regarding each tool, such that the end-user can make an informed choice about its appropriateness, use, and value.

Our second decision, and resultant guiding principle, involved the inclusion of assessment tools that are instructionally valuable. The first edition of this volume focused on program, student outcome, and basic skill assessment. For this edition (particularly in Part III), we include some assessment tools that have their primary purpose in classroom use. Many of the self-report instruments that assess trait or communicator style indicators are designed solely to give students feedback about their thoughts, attitudes, and self-perceived behaviors. They are not designed, nor should they be used, for any type of student outcome assessment. We include them in this volume, as there is no other publication that catalogues and describes these instruments. We believe that many communication instructors will find these instruments useful in their classrooms. To distinguish the two uses for the tools reviewed in Part III, that part is divided into two sections. The first section contains tools that can be used for skills and behavioral assessment. The second section contains instructional and informative instruments.

To support further the inclusion of various self-report measures, and measures of trait or style, we call the reader's attention to three main purposes or uses for the results of oral assessment: to advise and counsel students, to redirect curriculum content and pedagogy, and to respond to accountability concerns. The first purpose, advising students, can be achieved using self-report tools and measures of self-perceived traits or styles, such as those reviewed in Part III of this volume. However, we again caution the reader that these instruments were not designed for and should not be used for outcome assessment. For example, self-report tools assess only what a person knows or feels, not how that person might appear to an objective and trained evaluator.

Our third decision related to the inclusion of instruments where a paucity of information existed about the tool. For example, information related to psychometric concerns such as reliability, validity, or capacity to operate without biases could not always be located. But in some cases, the reviewers for Parts II and III are of the opinion that the instruments appear to be potentially valuable and worth inclusion. The reviewers indicate, on the review for each instrument, where such information is not available. The user of this volume should also be advised that the reviewers have not necessarily used all of the instruments reviewed herein. The intention of the review is therefore informative, not necessarily valuative. In some cases, information on "strengths and weaknesses" only addresses one of these two factors. If the reviewer has not had first-hand experience with the tool, comments regarding both strengths and weaknesses may be limited. Finally, the user should note that the prices indi-

cated on each review are the prices available at the time of this publication.

Our intention as editors is to make this second edition a valuable resource to many colleagues and administrators. Therefore, it also is our intent that this edition contain as much information as possible about potentially useful oral communication assessment tools. If we have overlooked a tool that should be reviewed and considered for inclusion, send it to us, presented in the review format used in this volume. We will have it reviewed and possibly added to the compendiums of tools presented here. Or, if you are aware of or have conducted tests of any tools already reviewed here, share that information with us. We invite your participation in editing a volume that will never be completed, that is an ongoing work!

A Preface Relating Alternative Assessment, Test Fairness, and Assessment Utility to Communication

Don L. Rubin, *University of Georgia*

The typical practice of educational and psychological measurement is guided by the adage, "If a trait exists, it exists to some degree. If it exists to some degree, it can be measured." The mission of the measurement enterprise, as traditionally construed, is encoded in the criteria of psychometric adequacy. That mission, stripped to its technical essence, is just this: Thou shalt maximize the proportion of trait-relevant variance in any test score.

Increasingly, however, the educational measurement community is coming to recognize that achieving accurate measurement is only one of several goals for assessment. Accurate measurement is necessary, but hardly sufficient. Some have argued that the notion of validity must expand to accommodate considerations of the *consequences* of assessment for learners, teachers, and communities (see review in Moss, 1992). A test that results in teachers devoting too much class time to rote learning or worksheet exercises may not be a valid test (Cronbach, 1988).

Given these more inclusive conceptions of assessment and measurement validity, what might be a proper mission for the community of communication educators and researchers who are concerned with the ways in which schools and colleges assess and test oral communication skills?

A decade ago—when the first incarnation of this volume was published—the mission was primarily to use assessment as a tool for consciousness raising. As a result of the Basic Skills legislation of the late 1970s (ESEA Title II), K-12 curriculum developers were charged, for perhaps the first time in recent U.S. educational history, to think about the career and life adjustment values of oral communication. Speaking and listening were to be treated as more than merely the precursors of literate behavior. At about the same time, blossoming interest in reforming the undergraduate experience in U.S. higher education focused attention on the role of oral discourse as a medium of learning and involvement across disciplines.

Those were admirable sentiments, but they were hardly tangible enough to drive the effort of *implementing* new speaking and listening curricula, for curricula are ultimately implemented not in school district offices or college faculty committees, but in classrooms. Assessment mandates can function as driving forces for implementing reform initiatives in classrooms. For example, teachers throughout the Commonwealth of Massachusetts began to plan seriously for oral communication instruction as a response to mandated formal testing of speaking and listening skills in the early 1980s. Teachers throughout Pennsylvania and Virginia followed suit in response to statewide initiatives for somewhat less-formal, classroom-based assessment methods.

Those efforts at raising educators' consciousness about the centrality of oral communication may be judged a qualified success. Witness, for example, the manner in which oral communication considerations have infused the proposed national standards for English and language arts (in draft form and under review as the present volume is being prepared for publication). At least one-quarter of the standards directly address classroom interaction and students' speaking and listening skills.

Having achieved this kind of positive impact during its initial stages, the movement for oral communication assessment might well look ahead toward pursuing additional missions. To achieve greatest mutual synergy, those new missions can profitably be located where well-established perspectives of communication arts and sciences intersect with emerging perspectives in measurement and evaluation. Accordingly, the following three missions are proposed for the second stage in communication assessment, the stage that this volume augurs:

◊ ensuring attention to developing and using alternative assessment methods, while continuing to adhere to appropriate standards for validity and reliability in traditional assessment;

◊ promoting test fairness by mitigating sources of cultural interference; and

◊ contributing to assessment utility by developing effective means for communicating results to stakeholders.

Alternative Assessment and Communication

The field of assessment is undergoing exciting ferment. Public schools as well as institutions of higher education are espousing goals such as "value-added education," "outcome-based education," and "national standards." Each of these goals is intimately bound up with evaluation. If we are to articulate the standards that we wish our students and teachers to exceed, then we must also devise the means by which we can recognize when those standards are met. We are entering an epoch in which assessment is at the forefront of educational reform.

At the same time, however, many educators are beginning to regard traditional psychometrically-driven testing as simply not up to the job of describing student growth and achievement in complex domains such as written expression or mathematical problem solving (Johnston, 1992; Wolf, Bixby, Glenn, & Gardner, 1991). Instead of standardized multiple-choice testing, the frontier of assessment is allied with practices like portfolio assessment, learning logs, and exhibitions. Herman, Aschbacher, and Winters (1992, p. 13) summarized these recent trends in assessment as involving five kinds of reforms:

1. Changes from behavioral to cognitive views of learning and assessment;
2. From paper-pencil to authentic assessment;
3. From single-occasion assessment to samples over time;
4. From single attribute to multidimensional assessment;
5. From near-exclusive emphasis on individual assessment to group assessment.

In virtually every case, alternative or "authentic" assessment programs involve performance-based evaluations. They may entail students directly demonstrating some task, or they may ask students to collect artifacts (in a portfolio) that indirectly document task performance. And these new types of assessment tasks ask learners to reflect on their performance and to communicate those reflections. Thus in a typical writing portfolio assessment, students choose some subset of the writings they have produced over a period of time. Then they may write a "biography" of each piece, explaining its genesis and the process by which they composed it. Or they may answer a series of questions inquiring why they regard each piece as important to their own development as writers (Camp, 1992).

Speech communication assessment has long been performance-based. It has considerable expertise to contribute to the present movement for alternative assessment. For example, public speaking assessment has continually wrestled with the issue of preparation for a performance test. Should students be provided speech topics ahead of time so that they can prepare in a realistic fashion, or do impromptu speaking tasks provide more accurate estimates of individuals' abilities? In addition, the experience of oral communication assessment can provide insight regarding ways to evaluate collaborative group processes, as well as optimal ways to conduct interviews about work products.

In exchange, the practice of oral communication assessment can be enhanced by adapting aspects of alternative assessment (Hay, 1994). Too many public speaking assessment procedures demand that students engage in "dummy run" extended monologues, rather than authentic communication. Few communication assessment procedures allow for sampling multiple performances across time; they rely instead on a single performance of unknown test-retest stability. Few oral communication assessment programs have been developed to exploit the power of portfolios to promote learner reflection. These remain as challenges for those involved in the assessment of oral communication competency.

Before departing from the topic of alternative assessment, it is worth remarking that the communication assessment movement need not—indeed must not—abandon the pursuit of rigorous psychometric adequacy for other kinds of measurement strategies. Though some proponents of alternative testing do sometimes seem philosophically opposed to reporting reliability coefficients, that ought not be the tack taken in the field of oral communication testing. Conventional indices of psychometric adequacy are discussed in Part I of this book, and it is crucial for the continued credibility of this enterprise that assessment specialists remain mindful of such indicators. There may be times, however (e.g., high stakes testing), when one sacrifices a degree of authenticity in exchange for a degree of reliability. Conversely, there may be times (e.g., program evaluation) when one trades off some internal consistency in exchange

for diffuse pedagogical impact. But these trade-offs need to be deliberate, and they ought to be considered short-term exigencies. Eventually we ought to be able to invent measures that are at once authentic and reliable. The use of a multiplicity of measurement approaches, some authentic and others more traditional, would yield a comprehensive assessment of oral communication.

Test Fairness and Communication

Typically, the process of test construction and of development—particularly for high stakes testing—includes several elements designed to minimize adverse impact on minority group test-takers. A bias review panel may review test items to eliminate culture-exclusive content. Statistical analysis will indicate items that function differentially for different minority groups, and such items will be eliminated from the item pool.

But what of the more subtle, and perhaps more pervasive, culture-typical communication demands of an assessment? Sociolinguistic test analysis (Rubin, 1992) is designed to evaluate ways in which the language and discourse patterns of assessment procedures can mismatch test-takers' own cultural norms about communication.

Three kinds of sociolinguistic interference can be detected. First, test-takers' oral dialect patterns may make it more difficult for them to ascertain the intent of a question. For example, a speaker of Black English Vernacular may experience some additional cognitive load in trying to decode a negative comparative question like, "Why didn't Jack close as many contracts as Mary?" Second, test-takers may have different discourse norms than those presupposed by a test. By way of illustration, members of certain Native American Indian groups might be very efficacious in presenting evidence in a persuasive appeal, but then consider it very rude to state explicitly for a listener the conclusion to which the evidence points. Finally, sociolinguistic analysis can point out that the entire testing situation might be more alien a communication context for some groups than for others. The idea of speaking with an interlocutor who sits stone-faced and silent while I respond to her question might be especially discomforting to people from some African American communities.

Test fairness in interview and communication assessment programs presents especially complex issues (Valadez, 1989). The very criteria by which most assessment instruments judge communication effectiveness—elaborated speech, unique expression, pointed argumentation, and the like—must be recognized as culture bound, characteristic mainly of middle-class Euro-American communication style (see reviews in Morreale, Gomez, & Shockley-Zalabak, 1992, and in Stiggins, Backlund, & Bridgeford, 1985). To hold minority test-takers to these standards may or may not be justified, but in any event must be understood as a statement about privileging certain culture stances and devaluing others.

A strong case can sometimes be made for a particular culture-linked communication practice as a *bone fide* job qualification. For example, the headquarters of a national organization might have legitimate warrant for wishing to employ telephone operators without marked regional accents. On the other hand, test-developers must be careful not to impose a culture-bound standard where none need be. Police dispatchers, for example, need only be highly intelligible. It is quite conceivable that a dispatcher could achieve that quality while yet retaining a speech style prominently marked for region or ethnicity.

Assessment Utility and Communication

An assessment procedure of the highest psychometric adequacy is of no value if its results are not effectively communicated to stakeholders (Cronbach, 1988). The stakeholders of educational assessment are easy to identify: students, parents, teachers, administrators, policy makers, and the public. Of these, the ones who often receive the least clear feedback from assessment—particularly from high stakes assessment—are students. Students of course are informed of the bottom line of assessment-based decisions: a passing grade, a failure to receive some certification, a placement in some particular program. Often, however, no one bothers to help students understand how *they* can use assessment data for decision making. (A marked exception to that generalization is the reporting of vocational aptitude assessment, which usually *is* deliberately geared toward client/learner decision making.)

Many school districts, it is true, make efforts to provide parents (if not students) with some background materials that will help them interpret their children's test results. But the "Standards for the Assessment of Reading and Writing" (Joint

Task Force on Assessment, in press) demand something further. They demand that results be reported to parents in a manner that will encourage dialogue between parents and children about the assessment.

Communicating assessment results for the purpose of promoting dialogue is exactly the point of "constructive evaluation" (Johnston, 1992). Constructive evaluation, with its concomitant dialogue among stakeholders, is most likely to ensue when assessors use examples of work products, teacher narratives, and other types of "thick description" to help report results. Assessment results that are adumbrated by thick descriptions of classroom discourse can empower classroom teachers to make their own local sense of the assessment results. Denuded numerical scores by themselves are not likely to support that kind of equity in utilizing assessment data.

Assessment can serve a crucial function in promoting the public dialogue regarding educational reform. Unfortunately, reports of test results are often followed in the public forum by sensationalistic indictments and accusations against educators, institutions, and students. This sort of sensationalism does little to promote productive public discussion. Were assessment results communicated along with more contextual information about the status of the schools and the nature of the assessment instruments themselves, the quality of the ensuing public responses would no doubt be more constructive (Joint Task Force on Assessment, in press).

These are some potential missions for communication assessment in the coming decade: to participate in alternative assessment, while continuing to utilize more traditional methods informedly; to work toward culturally fairer testing; and to promote the public dialogue regarding educational reform. These tasks build upon the work of the decade past; that is, they raise educators' consciousness regarding the centrality of oral communication.

To fulfill these missions, the most necessary ingredient is a critical mass of tests, measures, and assessment procedures that effectively measure various dimensions of oral communication. It is important that, collectively, these instruments exhibit diversity. They must gather data using diverse methods, and they must address diverse types of communication behaviors and skills. And it is important that these instruments exhibit accuracy in measurement. After all, we demand the highest quality of information to meet the challenge of school improvement.

The three-part present volume attests to the diversity of measurement instruments available within the domain of oral communication. It contains specifications that will permit test users to evaluate measurement accuracy. To compare the present volume with its precursor of a decade earlier is to appreciate the advances that have been made over that span. A decade hence, a similar comparison will reveal that communication assessment will have advanced by some additional quantum, and a new set of missions will emerge for communication assessment. And if we have labored well, we will have contributed to some progress in the teaching and acquisition of effective oral communication competency.

References

Camp, R. (1992). Portfolio reflections in middle and secondary school classrooms. In K. B. Yancy (Ed.), *Portfolios in the writing classroom* (pp. 61-79). Urbana, IL: National Council of Teachers of English.

Cronbach, L. J. (1988). Five perspectives on validity argument. In H. Wainer (Ed.), *Test validity* (pp. 3-17). Hillsdale, NJ: Erlbaum.

Hay, E. A. (1994). Alternative assessment in speech communication. *Journal of the Association for Communication Administration, 1*, (1-9).

Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Johnston, P. H. (1992). *Constructive evaluation of literate activity*. New York: Longman.

IRA/NCTE Joint Task Force on Assessment. (1994). *Standards for the assessment of reading and writing*. Newark, DE: International Reading Association.

Morreale, S. P., Gomez, A. M., & Shockley-Zalabak, P. S. (1992, November). *Oral communication competency assessment and cultural diversity: Ethnicity and students at-risk*. Paper presented at the meeting of the National Speech Communication Association, Chicago.

Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research, 62*, 229-258.

Rubin, D. L. (1992). Cultural bias undermines assessment. *Personnel Journal, 71*(5), 47-52.

Stiggins, R. J., Backlund, P. M., & Bridgeford, N. J. (1985). Avoiding bias in the assessment of communication skills. *Communication Education, 34*, 135-141.

Valadez, C. L. (1989). Problems of bias and test-wiseness in measuring oral communication and problem-solving skills through multiple choice items. In W. E. Tomes (Ed.), *Test-wiseness: Cultural orientation and job related factors in the design of multiple choice test questions* (pp. 17-27). Alexandria, VA: Assessment Council of the International Personnel Management Association.

Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of Research in Education, 17*, 31-74.

An Introduction to Assessing Oral Communication

Don L. Rubin, *University of Georgia*
Philip M. Backlund, *Central Washington University*
Sherwyn P. Morreale, *University of Colorado at Colorado Springs*

Why Teach Oral Communication Skills?

A well-known adage has it that of all the creatures inhabiting the earth, fish are the least likely to ever discover water. So it is with communication. Speech comes to us as part of our innate endowment as human beings. We are engulfed by communication in all our daily affairs, and usually we are not directly aware of our oral communication environment. But it is nonetheless vital to our well-being and survival.

Competencies in oral communication, in speaking and listening, are prerequisites to success in life and in school. Most instructions for classroom procedures are delivered orally by teachers. Consequently, students with ineffective listening patterns fail to absorb much of the material to which they are exposed. Their problems are intensified when they respond incorrectly because they do not decode questions carefully. Students who listen poorly are often isolated and left out of classroom activities. Speech production also affects academic achievement. Students who cannot articulate their knowledge are judged ignorant. Additionally, some speech styles trigger stereotyped expectations of poor ability, expectations that are likely to be self-fulfilling. Quiet children may be appreciated for their "good behavior," but they are subject to similarly negative school experiences (McCroskey & Daly, 1976). Students who do not ask for assistance will not receive it, and research typically indicates that reticent students progress more slowly despite normal levels of aptitude.

Beyond the confines of school, oral communication competency can contribute to social adjustment, satisfying interpersonal relationships, and, ultimately, professional success. Youngsters with poor communication skills are sometimes viewed as unattractive by their peers and enjoy few friendship bonds (Hurt & Preiss, 1978). Antisocial and violent behavior often go hand-in-hand with underdeveloped social sensitivity and lack of conflict resolution techniques. Remediation programs have reduced the incidence of antisocial acts by means of communication training, and counselors acknowledge that many family problems are caused by poor communication that may be ameliorated by improving communication between family members (Galvin & Brommel, 1991).

Communication competence is no less crucial in the marketplace and in professions. Communication skills rank high among lists of managerial competencies and prerequisites for success in the workplace (SCANS Report, 1993). Professionals—doctors, lawyers, engineers, teachers—require more than just subject matter expertise. These professionals must listen effectively to their patients, clients, or students in order to identify and analyze problems. They must speak effectively to implement their solutions. And even blue collar workers have frequent occasion to engage in job-related speaking and listening. In summary, empirical research has consistently related oral competency and communication training to professional and academic success (see, for example, Curtis, Winsor, & Stephens, 1989; Rubin & Graham, 1988; Rubin, Graham, & Mignerey, 1990; Vangelisti & Daly, 1989).

On another note, speech curricula have traditionally stressed the importance of communication for the preservation of a democratic society. Throughout its history, the United States has fought vigorously to safeguard freedom of expression under the assumption that full citizen participation is the surest guarantee against tyranny. Not every citizen deliberates as a member of a legislative body, but numerous opportunities for citizen input are available. At the very least, citizens are responsible for staying informed, and much of the pertinent information must be culled by listening and utilizing one's communication abilities.

Finally, oral communication is essential to full psychological development. Early research called our attention to the fact that self-concept is acquired through interaction with others (G. H. Mead, 1934). Additionally, self-actualization, a sense of fulfillment (Maslow, 1954), usually involves interpersonal activities: making contribu-

tions, exerting influence, or being recognized in a social manner. In addition, communication is a means for artistic expression as well as self-discovery.

The fact that all students come to school with some basic speaking and listening skills and also seem to develop more mature communication behaviors on their own as they grow older does not imply that all students are *effective* or *competent* communicators. Educators occasionally comment, "My students don't need to learn how to talk. That's one thing they do too much of already." But competent communication must be cultivated. Students who communicate well in familiar settings may lack the confidence and flexibility needed to express themselves effectively in a wider range of situations. Educators cannot rely on haphazard, unguided learning outside of the classroom to impart communication effectiveness. Systematic instruction in oral communication is imperative.

Yet of all the basic competencies, speaking and listening are still frequently neglected in many schools. This neglect is all too typical despite recent curriculum reforms that call attention to the importance of oral abilities and the need to implement oral communication instruction (Berko, 1994). Teachers are being held accountable for students' performance on many mandated tasks. But teachers are generally not held accountable for teaching students to communicate with competence. Furthermore, few teachers, except for those whose degrees are in communication, have received training in communication education or have materials available to aid instruction. Consequently, little concerted instruction in speech communication takes place in kindergarten to grade 12, whereas in many 2- and 4-year colleges and universities, speech communication and its instruction are beginning to receive considerably more attention.

Oral Assessment for Classroom Decision-Making

If students' speaking and listening proficiencies were systematically evaluated, it is likely that schools would begin to implement more oral communication instruction. One substantial benefit of large scale assessment of oral communication is that such testing can guide innovation in the curriculum. Indeed, early experience in Great Britain and elsewhere demonstrates that speech assessment has a "washback" effect on the amount and kind of speech teaching undertaken in classrooms (Barnes, 1980). Another benefit of oral communication assessment is that test results can be used to make decisions about the best manner in which to place individual students in instructional sequences. Assessment procedures that yield micro, rather than macro, judgments can be used for individual diagnostic purposes (D. L. Rubin, 1981). For example, students who are assessed as having difficulty with fundamental vocal production factors might concentrate on oral reading and presenting, while those whose difficulties lie in the area of organization might focus on developing skills related to organizing and outlining. Students who demonstrate strengths in, say, literal comprehension of spoken materials might advance to instructional units emphasizing more critical listening.

Oral Assessment for Programmatic Decision-Making

Speaking and listening tests can also provide valuable information for program evaluation. Because large scale programs of oral communication improvement are new to some academic institutions, it is especially important to evaluate their effectiveness and to secure data that will enable these programs to be "fine tuned." Program (and teacher) effectiveness is best judged with reference to student achievement on program objectives. If students are not achieving criterion performance levels in language use, for example, teachers and administrators will recognize that additional instructional effort needs to be directed to this area. It is worth noting, however, that student achievement can be interpreted as an indicator of program success only when student aptitude and institutional resources—the raw materials with which the program has to work—are also taken into account. Also, student achievement is not the only data that might contribute to program evaluation. Attitudinal outcomes, self- and peer-evaluations, are also useful information for this purpose. A recent examination of exemplary models of program assessment in oral communication can inform this process for the teacher or administrator (Hay, 1992).

Oral Assessment for Certifying Student Competencies

Another use for speaking and listening assessment tests is to certify students as having attained

(or not attained) levels or degrees of oral communication competency. Certification in basic skills may be increasingly demanded by standards-based education movements such as Goals 2000 (National Education Goals Panel, 1992). Surveys of state and local jurisdictions indicate that many have already adopted large scale tests of speaking and listening skills as a means of certifying students' competencies in basic skills (Litterst, VanRheenen, & Casmir, 1994).

In summary, the process and results of assessment of oral communication can be of benefit to students, teachers, and administrators. For students, assessment results can be used for advising, counseling, and placement purposes in addition to ensuring that students have developed the expected and necessary oral communication competencies prior to graduation. For faculty, examination of assessment results can be used to revise course content and to redirect teaching. For administrators, assessment of oral communication can be used to improve curriculum structure and resource allocation, and to assure accreditation agencies and the public that each college graduate has acquired the necessary knowledge, skills, and behaviors to communicate competently.

Background to This Publication

Developments in the past decade motivated the present revision publication and its precursor. With the 1980s background of a nationwide movement toward competency-based education, speech communication educators recognized that their field was not exempt from the challenge of accounting for educational outcomes. An early step in facilitating that accountability was the publication of the results of the National Speech Communication Competencies Project (Allen & Brown, 1976). That document examined the development of communication skills and promulgated a description of functional communication competence that has helped guide many subsequent efforts in curriculum, instruction, and evaluation. Then, in 1978, SCA published *Assessing Functional Communication* (Larson, Backlund, Redmond, & Barbour) as the first description and review of available assessment instruments for communication competencies. The development of lists of competencies in speaking and listening for high school graduates (Bassett, Whittington, & Staton-Spicer, 1978), sixth-grade students (Backlund, 1985), and community college students (SCA, 1990) has likewise influenced teaching and assessment. In their "Standards for Effective Oral Communication Programs," the American Speech-Language-Hearing Association and the Speech Communication Association (1979) asserted that effective instructional efforts must include provisions for appropriate and constructive methods of assessment and evaluation. Such methods were further clarified by the development of the "Criteria for Evaluating Instruments and Procedures for Assessing Speaking and Listening" (SCA, 1979) that outlined standards by which oral communication assessment instruments and procedures would be judged. These criteria were updated and refined in 1992 when SCA developed, and then the following year published, "Criteria for the Assessment of Oral Communication: A National Context" (SCA, 1993). The Speech Communication Association has also published sources of instrument development information and support such as *Oral Communication Assessment Procedures and Instrument Development in Higher Education* (Rubin, Moore, Sisco, & Quianthy, 1983) and *Resources for Assessment in Communication* (SCA, 1985). A national conference on assessment issues was held in the summer of 1990 in Denver, Colorado, sponsored by the Committee on Assessment and Testing, the Educational Policies Board, and SCA. Finally, a summer conference on assessing college students' oral competencies was convened in Alexandria, Virginia, in 1994. That conference addressed an array of issues and concerns related to assessing undergraduates' oral competence. The compiled proceedings of the conference are available through SCA (Morreale, Berko, Brooks, & Cooke, 1994).

These efforts to understand communication competency produced practical knowledge about the various ways to assess oral communication. For example, R. B. Rubin's (1982) *Communication Competency Assessment Instrument* has been used for over 10 years to measure overall communication competency. In 1993, Morreale, Moore, Taylor, Surges-Tatum, and Hulbert-Johnson developed and tested *The Competent Speaker* as a tool to assess public speaking behaviors. In 1995, Spitzberg completed testing and published *The Conversational Skills Rating Scale*. In general, however, assessment programs in public school and in higher education have been limited by some lack of awareness of the availability of suitable instruments for assessing oral communication. Consequently, the Committee on Assessment and Testing and the Communication Assessment Commission of the Speech Communication Association are acting, through this volume, to identify existing

instruments and further the development of additional instruments for the measurement of communication in a variety of contexts. The present volume represents an effort to address a perceived paucity of oral assessment tools, despite the prior publication of similar compendiums of instruments (Rubin, 1994; Spitzberg & Cupach, 1989).

Objectives of This Publication

Four primary objectives guided the development of the present publication:

◊ to identify and examine existing assessment instruments in oral communication for two academic levels: kindergarten through grade 12, and colleges and universities;

◊ to abstract and describe assessment instruments and systematically report their availability to interested teachers, scholars, and administrators;

◊ to provide background for research that is needed to develop oral communication assessment instruments for elementary, secondary, and postsecondary schools; and

◊ to encourage the development of new instruments and the testing and utilization of existing instruments of merit.

For this publication, SCA's (1993) "The Speech Communication Association's Criteria for the Assessment of Oral Communication" was used to review instruments for inclusion in this volume. Particular attention was paid to assessment instruments that evaluate communication behaviors, as opposed to instruments that describe behaviors but assign no judgments of quality. In addition, emphasis was placed on measures of communication per se, verbal and nonverbal encoding and decoding in situations ranging from high interaction (interpersonal communication) to extended and uninterrupted discourse (public speaking). The effort was directed toward instruments whose main purpose was the measurement of the behavioral or affective domains of communication, or the measurement of a related construct, such as conflict style, etc. The effort did not extend to instruments that directly measure knowledge or cognition in regard to communication. However, inferences about cognition (and affect) can be made based on assessing observed communication behaviors. Finally, attempts were made to include instruments appropriate for a variety of individuals including nonnative speakers, students of culturally diverse backgrounds, and students with special needs.

Procedures for Reviewing Assessment Instruments

To begin the revision of this volume, committee members volunteered to research selected areas of instruments. The reviewers extensively surveyed the available sources, compared instruments to the SCA criteria, and developed reviews of each instrument according to a common format. A catalogue of instruments that met the criteria of the reviewing committees is presented in two tables. A table and reviews of instruments for grades K-12 appear in Part II of this publication. Two tables corresponding to two sections of reviews for instruments for 2- and 4-year colleges and universities appear in Part III. The first table and section contain skills/behavioral assessment tools; the second table and section contain instructional/informative instruments. The contents of the reviews reflect the views of the individual reviewers as influenced by their expert judgment.

In reviewing the instruments, two points provided general guidance to the reviewers. First, in the absence of a consensually acceptable model of competent communication, it is difficult to evaluate instruments' construct and content validity. Objectives and competency lists adopted by one assessor may diverge widely from those that guide test selection by another assessor. Second, it is anticipated that some users of these reviews will not be speech communication scholars, but rather evaluation specialists and school administrators. The instrument review form reflects the general concerns of this target audience with respect to psychometric adequacy and administrative feasibility. More specific discussions of the review and selection process for the instruments are included in the introductions to Parts II and III. To facilitate the reader's use of Parts II and III, the concepts of validity and reliability are reviewed here as they relate to the instruments reviewed in both Parts II and III.

Instrument Validity

The section of the instrument review form on validity is concerned with the extent to which the instrument actually measures the skills or knowledge it intends to measure. Validity may be determined in many ways, and the presence of multiple

validity studies using different methods and different target populations strengthens the case that the instrument actually measures what it purports to measure.

Predictive validity deals with the ability of the instrument to predict performance on another measure that is known to be valid and that is theoretically related to the instrument in question. For example, a test of communication competence might be assumed to predict success in jobs that rely heavily on oral communication.

Concurrent validity is similar to predictive validity except that it focuses on the relationship between individuals' performance on the instrument in question and on other instruments that measure the same thing. If a group of students were administered a speaking performance test and were also rated by their speech teacher, then the correlation on these two measures would be a test of concurrent validity of the speaking performance test. Or, if students were administered two different listening tools, one already tested would support the validity of the other. Both predictive and concurrent validity are referred to as criterion validity.

Content validity indicates the degree to which the content of an instrument represents the domain of knowledge and skills it intends to measure. Content validity is usually determined through expert judgment. One common method is where experts are given a description of the test's objectives and then asked to categorize each item by these objectives. Content validity is measured by the degree of agreement among judges in the category assignments. Content validity is also referred to as face validity or intrinsic representativeness.

Building a case for *construct validity* takes many forms. Construct validity may derive from any experiment that sheds light on the nature of the phenomenon that the instrument is trying to measure. Common bases for construct validity include factor analysis of the items an instrument contains and the known groups method. Also, the instrument has construct validity if it can be theoretically supported in terms of its method of development or base in the literature.

Instrument Reliability

The section of the review form on reliability reports the measurement accuracy of the instrument. There are various methods for determining reliability.

Test-retest reliability measures the stability of an instrument over time. Assuming that the respondents have not been exposed to instruction and have not undergone a major growth in the knowledge or skills being measured, they should receive approximately the same score on an instrument at two points in time. However, test-retest is not always a good measure of reliability, especially of skills that might be expected to change.

In some cases instruments are designed to have alternate forms that are equivalent in content and difficulty. The correlation between individuals' scores on the different forms is a test of *alternate forms reliability*. When we offer students a choice of topics about which to speak in a speech performance evaluation situation, we are assuming that these topics constitute equivalent forms of the same test. Similarly, when we use different small group configurations or different interview scripts to elicit oral performance, we must assume that these constitute equivalent forms.

Taking the concept of alternate forms reliability a step further, it is possible to think of an instrument as a random set of items, each of which is a "test" of some part of the content domain. The degree to which the respondents' performance on one item is related to their performance on other items is a measure of *internal consistency reliability*. Typically, a correlation coefficient is calculated to analyze the relatedness of items. For example, a Cronbach's alpha considers the correlation of each item and all other items on the instrument, but the Cronbach's alpha is not meaningful if the instrument in multidimensional.

Tests of performance are markedly different from paper-and-pencil tests. For these tests, measurement takes place within the person who assigns the rating or score. Here the reliability of the scorer is at issue, not the reliability of the test. *Scoring (or interrater) reliability* can be assessed by having more than one person rate the same performance. The correlations or percentages of agreement between raters in these ratings is a test of scoring reliability. Usually scorers are evaluated for reliability after training but before they begin rating. However, to insure that scorers remain consistent over time, it is important to check their reliability periodically during the scoring process as well. Additionally, one rater using a valid instrument can be trained to use that instrument reliably, eliminating the need for multiple raters.

As a part of the development of some large scale assessment instruments, *norming* or *criterion referencing* studies are conducted, and these are also discussed on the review form. For these studies, the instrument is administered to a large number of respondents, and the results provide performance benchmarks for future users of the instrument. Norming studies for standardized achievement tests yield charts that transform raw scores into normed scores, most frequently grade equivalence. Standard setting studies are sometimes conducted for tests that measure mastery of specific objectives—that is, criterion referenced tests. Data collected from samples of students are usually compared with data from another source, such as teacher ratings, to determine what test scores represent mastery level. A caution for all norming and criterion referencing data is that the characteristics of the original population assessed may be different from the population that the present user is assessing.

Issues In Assessing Oral Communication

Major issues related to the use of oral communication assessment instruments should be considered by the user of any instrument. Several of the most important of these issues are described here to assist the reader in determining the best procedure available to meet the goals of his or her assessment program.

Content of Assessment Instruments for Speaking and Listening

N. A. Mead (1982) makes a useful distinction between process knowledge and content knowledge as related to the goal of assessment. For most academic areas, the educational experience involves acquiring knowledge for application later in life. Testing in those areas (such as history, the sciences, etc.) focuses on the student's level of knowledge acquisition. However, oral communication is a process skill, and, as such, it makes more sense to assess the student's performance with speaking and listening than it does to assess the student's knowledge about speaking and listening. Thus, virtually none of the instruments reviewed in this volume assesses the student's knowledge of the communication process. The content of the instruments focuses on skill acquisition and attitude development. Educators usually are more interested in what students *do* with speaking and listening skills rather than in how much students *know* about those skills. Comments about content below concern themselves with student performance and skill in regard to speaking and listening, respectively.

Speaking. The content of speaking assessment procedures is varied. One way in which this content can be categorized is in terms of mode of discourse. At the elementary age level, most speaking tasks are either narrative or descriptive. A number of tests designed for nonnative speakers also rely on storytelling. For older native English speakers, greater variety is evident. The speaking tasks often call for exposition in the form of extended monologues. Other modes of discourse include extended persuasive monologues or simulated persuasive conversations, telephone conversations, introductions, and responses to questions in an interview.

Speech assessment procedures can be categorized in terms of communication situations as well. In particular, it is useful to examine the types of audiences that are featured in oral performance tasks. Of course, in all speaking tests students will be aware of the examiner as an ultimate audience. However, in the majority of instruments reviewed, the examiner is the sole audience to whom students speak. It should be noted that speakers do not typically communicate in order that their oral proficiency be evaluated. Indeed, evaluation usually inhibits communication. To the extent that assessment procedures offer no pretense for speaking other than evaluation, these procedures may yield somewhat inaccurate samples of communication performance.

A single examiner-audience is most natural in interview situations. One pitfall of interview situations is that the interviewer may exert overriding influence on students' speech behaviors, resulting in considerable unreliability. A single examiner-audience is most anomalous in those situations in which students are called upon to deliver a speech to that individual; we do not usually deliver formal speeches to an audience of one. However, this type of testing situation has the advantage of not being confounded by audience reaction. The problem of unnatural audiences is somewhat relieved by procedures that simulate situations involving realistic speaker/audience relations. These procedures may ask students to simulate an emergency telephone call to a police operator, give directions to a stranger, or persuade a friend to grant some favor. These simulation tasks, however, may confuse

speech proficiency with role-playing ability. A balance must be struck between the conflicting needs of internal validity (testing what you think you are testing) and external generalizability (to real situations).

Listening. The content of listening assessment instruments and procedures is as varied as that of speaking tests. Listening is not a unitary skill, but it is rather a complex of subskills, each of which is brought into play to greater or lesser degrees depending on the nature of the listening task (Lundsteen, 1979). It is natural, therefore, that tests of listening ability tap a variety of skills. Test users should make sure that the selected listening test conforms to their particular measurement objectives.

Most often, listening tests measure literal comprehension of spoken material. It should be noted that comprehension is generally confused with recall or retention, because questions typically follow some extended discourse. Some testing methods alleviate this confusion by providing a verbal context that may lessen reliance on memory. Similarly, tests that deliberately select brief passages and present few questions for each passage may tax memory to a lesser extent.

Many listening tests focus on listening for directions, a type of purposeful listening that is readily measured by accuracy of behavioral response (e.g., circling the correct item, drawing the proper path on a map). Other listening skills frequently measured include recognition of speaker's purpose, or making inferences or interpretations beyond material given. Other listening instruments include subtests reflecting ability to interpret paralinguistic cues and also ability to render social judgments from speech. Some subskills such as vocabulary, syntax, and phoneme recognition are critical to communicative listening but are so narrow that they might better be considered receptive language or decoding skills. Phonemic discrimination and identification are viewed as essential for reading readiness but should not be construed as measures of listening ability.

Outside of the classroom, the bulk of listening activity takes place in the course of interaction. When referential communication tasks assessing skills of both the speakers and the listeners permit free oral interchange, these tests approximate interactive communication. Other instruments measure interactive listening skill more indirectly by including conversational speech among their listening passages.

Criteria of Assessment Instruments

Another issue related to the content of assessment instruments that requires examination is their evaluation criteria. Along what dimensions of quality do the instruments render judgments? Listening tests are primarily concerned with accuracy of recall, of following directions, and of perceptions about social relationships. Conceivably, tests could be devised that provide information about listening activity as well. Such instruments would indicate what type of listening (critical, aesthetic, informative) students are engaging in during the course of a stimulus passage, and the degree of concentration or constructive assimilation that characterizes their listening processes.

Speech assessment procedures exhibit a fair amount of consistency in their evaluation criteria. Typically speech rating scales reflect only three clusters of judgment, despite the fact that they may include a larger number of variously labeled criteria. These clusters typically are content, delivery, and language. These criteria, with the addition of organization, account for most speech performance rating scales reviewed in this publication. Despite similarities in criteria, rating scales differ in the weight accorded each criterion. For example, instruments may vary in their treatment of language. Some instruments may weight language most heavily of the criteria, whereas others apportion emphasis more equally among all dimensions of speech evaluation. The definition of language quality adopted by some instruments stresses conformity to the conventions of standard American English. Other instruments, particularly those designed for nonnative speakers, convey more detailed information about the types of grammatical structures mastered.

Just as some listening tests were characterized as narrow tests of receptive language, so some speaking tests are measures of productive language, not of communication. This is certainly true of procedures that ask students to imitate words or sentences in isolation, and then apply criteria that evaluate articulation or grammatical interference of a first language. It is no less true of procedures that incorporate some communicative context like an interview, and then rate speakers on exclusively linguistic grounds. Merely eliciting language by means of a communicative task does not constitute a test of communication competence.

A few speaking instruments that emphasize language quality criteria reflect the contextual and interactive aspects of communication better than

many of the more conventional rating scales. These instruments measure the degree to which language is appropriate for or adapted to the demands of the communication task. For example, ratings of a response may depend on the type of question asked. Or a test may measure the degree of elaboration, not just simple labeling, that is expected in a response to a narrative task. Rating scale items may express communication-oriented criteria like "appropriateness" or "intelligibility" rather than formal linguistic properties like sentence structure, standard usage, or correct pronunciation.

Responses and Scoring Procedures Used in Assessment Instruments

Multiple-choice formats are the stock-in-trade of standardized testing. Questions are designed so that each has a single correct answer; tests can be graded easily by machine or template without any problems of unreliability in scoring. Item difficulty is readily ascertained and controlled, and test forms can be equated by well-established methods. Some indirect tests of speaking ability attempt to utilize multiple-choice responses, but the technique is more widely represented among tests of listening proficiency. Multiple-choice questions can be used not only to measure literal comprehension but also to assess higher order abilities like recognition of speaker's purpose and inference-making, and aspects of critical listening. One of the drawbacks of many multiple-choice listening tests, however, is that students must read printed questions and response alternatives, thus confounding listening ability with reading ability. Some listening tests combat this problem by using tape-recorded presentations of questions and response options. Others use pictures instead of verbal response options. Another technique employed in some measures of listening skill is behavioral response. In particular, this type of performance measure is used in tasks involving directions. In general, these tasks approximate normal listening activity, and thus they are more valid than less direct measures.

The most common means for assessing speaking skills are performance rating scales. D. L. Rubin (1981) and Stiggins, Backlund, and Bridgeford (1985) discussed a number of factors pertaining to the use of this technique in large scale assessment. Their major disadvantages lie in the potential for unreliable scoring and in the relatively large expenditures of staff time. Some systems seek to avoid the costs of external raters by having classroom teachers evaluate students' typical or elicited speech. This approach would seem to exacerbate the problem of rating error, and it begs the question of time allocations. Other alternatives to using performance rating scales in assessing speaking ability are techniques that use particular discourse features as indicators of quality of expression. These alternatives include the use of measures of syntactic complexity for assessing oral proficiency and the use of total number of words uttered, lexical diversity, articulation, and sentence expansion to measure linguistic features.

When speaking tasks are structured in a way that permits objective measurement of success, it is possible to derive measures of communication effectiveness. For example, it is possible to use "shift of opinion ballots," which ask audience members to indicate their attitudes toward a topic both before and after the delivery of a persuasive speech, to measure the effectiveness of the speaker. Referential communication tasks (Dickson & Patterson, 1979) measure communication effectiveness of a speaker by seeing whether a listener can identify the correct object from an array based on the speaker's description of the object. Effectiveness of small group communication can be evaluated by assigning a problem-solving task to a group and then recording the accuracy and speed with which the group solves the problem. These techniques, however, do not elicit uncontaminated measures of individual communication competence, because audience characteristics, listener skill, and group composition are factors beyond the control of the speaker and can affect communication success. The effectiveness of some referential communication tasks, however, can be assessed without recourse to measuring listener accuracy. For example, some tasks require the speaker to state the attributes of an object or geometric figure that uniquely describe it. Communication effectiveness is evaluated simply by counting the number of critical features that the speaker identifies (Piche, Rubin, & Turner, 1980).

Administrative Feasibility of Assessment Instruments

If measures of oral communication competency are to be adopted for large scale assessment programs, they must be administratively feasible. Such feasibility considers issues related to amount of student time, reasonable allocation of personnel for administration and scoring, and the need for specialized training for administration, scoring,

and interpretation. Communication is a complex, interactive behavior. And oral communication is a dynamic and ephemeral process, leaving behind no permanent trace. Therefore, tests of communication competence, speaking particularly, are apt to be more expensive in terms of labor and money than many other large scale assessment procedures. Administrative support for such assessment procedures could be an indicator of those administrators' commitment to students' oral competency.

Many tests of listening ability are, however, amenable to group administration. Even skill at following directions can be assessed in this manner. Tape-recorded administration instructions and response options not only reduce unreliability and confounding with reading ability, but they also contribute to ease of testing. Some listening measures, on the other hand, allow for a wider range of response modes, and these require individual administration.

Tests of speaking skill conducted as interviews or as extended monologues naturally demand individual administration. Moreover, it is advisable to use multiple raters to insure reliability. This can be accomplished by assigning two staff members for "live" rating or by tape-recording performances for subsequent evaluation by two raters. One other approach attempts to reduce the testing burden by requiring classroom teachers to screen their students based on their typical classroom communication behavior and to refer only those students "in question" for individual assessment. However, there is some evidence that suggests such screened ratings are subject to bias and are not reliable. As already suggested, another approach is to train a group of raters to a high level of reliability and allow them to rate individually.

It is possible to reduce administration costs by using group communication tasks, since a number of students can be evaluated during the course of, say, a 20-minute discussion session. Similarly, referential communication tasks may also be adapted to simultaneous administration to several dyads. The least practical but perhaps the most valuable methods of oral examination are those that require transcription and analysis of speech samples by multiple raters. Summarily, feasibility issues should involve consideration of the SCA criteria that suggest that for oral communication skills to be assessed, the student must either speak or listen (SCA, 1993).

Target Populations and Potential Sources of Bias

The instruments reviewed in Part II are intended for the K- 12 grade range, although the elementary grades receive particular emphasis, especially among commercially-developed instruments. The instruments reviewed in Part III are intended for all ages at two- and four-year academic institutions, and also include instruments categorized as being for adults. In choosing among these instruments, the user should be aware of potential biases in the testing procedure. Instruments vary considerably in their efforts to minimize group bias effects. Some technical manuals document the work of minority group reviewers who examined items in order to eliminate potential bias. It should be noted, however, that differences in central tendency are not, themselves, evidence of test bias. Rather, a test is biased if it over- or under-predicts scores on some independently administered criterion measure (Cleary, 1968). If an instrument contained in Part II or Part III has been tested in any way for any form of bias, that test is reported in the review. However, any identified biases may be related to cultural or gender differences, as opposed to biases in the instrument itself. That is, the instrument may identify an actual difference in behaviors, or self-perceptions related to the subject's cultural background or gender. In the final analysis, the user or assessor should be sensitive to whether an instrument operates without bias relative to the particular target population being assessed.

Locally Developed Assessment Instruments

Developing instruments locally for assessing any dimension of oral communication requires considerable time and effort as well as familiarity with measurement and content concerns. Often it is not feasible to submit locally developed instruments to the same degree of technical review for reliability or validity as commercially developed instruments. However, there are some situations where local development of assessment instruments is desirable. For example, a school district may adopt a set of specific speaking and listening competencies and develop an instructional program directed toward building those competencies. In order to measure its success, the district may find that it is better to develop a test locally that is tailored to its specific competencies than to use existing tests that only measure some of those

competencies or that only measure those competencies indirectly. The following brief step-by-step descriptions of the development process provide direction to local agencies that wish to develop their own instruments.

Speaking

In determining the speaking skills and types of speaking tasks that are important, local developers must first define the types of speaking skills and tasks students should be able to perform. In developing this list, developers will find it helpful to review curricular objectives, instructional materials, and teaching practices. They should involve a full range of people concerned with the results of the assessment process, (e.g., teachers, curriculum specialists, administrators, parents, and students). The resulting list may focus on specific skills important in all speaking situations (e.g., speaking distinctly or speaking in an organized fashion). The list may also focus on specific tasks that are considered important (e.g., giving directions, giving a speech, or asking questions). It is critical that the speaking skills and tasks listed be as specific as possible, so that they may be observed and measured.

Two types of approaches are used in assessing speaking behaviors. First, in an observational approach, the student's behavior may be observed and assessed unobtrusively. Second, in a structured assessment approach, the student may be asked to perform one or more structured speaking tasks, and his or her performance on the tasks is then assessed.

If an observational approach is taken, the developer must decide what speaking behaviors will be observed (e.g., asking questions, responding to questions, or speaking in group discussions). In addition, the developer needs to decide how many times each student will be observed, and for how long. The observer may be the regular classroom teacher, or someone from outside the classroom, such as a teacher from another grade level, a chairperson, or a counselor. It should be noted that optimal evaluation can only be provided by a well-trained speech communication professional.

If a structured approach is adopted, the developer must decide what type of speaking tasks will be used. Also, the developer must decide on the setting for the tasks. The student might be asked to perform certain tasks in front of the entire class, in a small group setting, or in a one-on-one situation with the assessor. Again, the assessor may be the classroom teacher or someone from outside the classroom. But again, training in speech communication will contribute to optimal evaluation.

Next, a scoring system that describes acceptable and unacceptable levels of performance for the speaking skills or tasks already identified in the first step must be developed. The scoring system may involve a two-point determination (e.g., the behavior of interest is either present or absent, or the student can be heard or cannot be heard). Alternatively, the scoring system may define a continuum of behaviors that ranges from lowest to highest (e.g., while speaking, the student is very disorganized, somewhat disorganized, fairly organized, or very organized. However, when a continuum is used, it is necessary to describe each level of the scale in terms of specific behaviors that represent that point in the scale. The resulting scoring system will be used either for observation ratings or structured ratings, as determined previously in the second step.

Once the basic approach is established and the scoring system is developed, it is necessary to train raters in the use of the system. Training should include thorough instruction in the categories in the scoring system, provision of examples of performance that represent the various categories, and opportunities for the raters to practice rating student performance. Raters should have ample opportunity to ask questions about the categories and discuss their practice ratings.

Often the training will lead to alterations in the scoring system. It is possible that some initial distinctions made in the scoring system will prove impossible to observe in actual performance. Once the system is finalized and raters are comfortable in their abilities to make ratings, raters should be tested for interrater reliability. They should be given several samples of performance and asked to rate them without discussion. The degree of agreement among the raters is a measure of interrater reliability. Raters should also be trained in test administration procedures.

The final steps in the assessment process are data collection and scoring. These activities may happen simultaneously or in stages. Ratings may be made on the spot, or the speaking performance of students may be audiotaped or videotaped and scored at a later time. The advantage of recording performance is that it allows for scoring in a more controlled environment.

In addition to testing interrater reliability at the end of training, the reliability of the ratings also should be checked during the assessment period. While effectively trained raters can conduct reliable assessments individually, periodic multiple rater assessment should be used to check interrater reliability. Checking the reliability does not have to occur for every rating, but it should be conducted at random for a reasonable percentage of the ratings.

Listening

To determine the knowledge of the listening process, the listening skills, and types of listening tasks that are important, local developers may begin by defining the types of listening knowledge the students should know and the types of listening skills and tasks that students should be able to perform. The steps in this process are the same as they are for speaking. The resulting list may focus on skills that are important to all listening situations (e.g., understanding main ideas and details). These skills may be similar to reading comprehension skills. The list may also focus on specific listening tasks that are considered important (e.g., listening to directions, listening to a lecture, or listening on the telephone). As with speaking, it is critical that the knowledge, skills, and tasks listed be as specific as possible so that they may be objectively measured.

The next step in developing listening assessment instruments is to assemble stimuli that the students will listen to in the assessment. These stimuli should reflect the listening tasks identified in the first step. Listening material may be drawn from existing sources. Natural listening material such as public service announcements, commercials, or news stories make particularly good material. It is also possible to write material that particularly reflects the tasks identified in the first step. Care should be taken to use material that is relatively short, is interesting to students, and does not reflect a bias toward a particular gender, racial/ethnic, socioeconomic, or geographic group.

The actual production of stimulus material may take two forms. The material may be written in script form so that it may be read aloud by the test administrator, or it may be recorded on audiotape or videotape. The advantage of taped materials is that they guarantee standard administration and allow for variety in stimulus material, such as various voices, conversations, or sound effects.

Several possible types of listening items may be developed. The most typical is multiple-choice items that ask a question about the listening stimulus and provide several possible response options. Another type is short-answer items that ask a question and require the student to write or give orally a short response. A third type, used for following direction tasks, presents graphic material, such as a map, and asks the student to complete a certain task, such as drawing a route onto the map. A variation of this listening item is to describe an object and ask the student to draw the object or to select the appropriate object from a set of pictures. In all cases, item development should follow established item construction standards that may be found in measurement textbooks.

It is impossible to identify all the possible confusing or problematic aspects of stimulus materials or items until they have been field tested with a sample of students who are similar to those who will be assessed. The results of field testing may be used to pick the best stimuli and items. Measurement textbooks provide some simple techniques for reviewing field test data. In addition, field testing provides information about the amount of time it takes most students to complete the items.

This information should be used to establish the time limits for the finalized test.

What Should Be Tested? Research and Development Priorities

Ideally we should test what we teach or what we expect to impact. But in many cases, adoption of standards for oral communication competency by state and local educational agencies is a new phenomenon. Moreover, even when such competencies are mandated, the extent and fidelity with which they filter down to classroom practice is unknown. Therefore, a priority in assessment is to continue efforts to determine current classroom practices in speech communication. The Committee on Assessment and Testing sponsored a number of such survey efforts in the 1980s, beginning with that by Backlund, Booth, Moore, Parks, and VanRheenen (1982). Those survey efforts have been updated, with the most recent occurring in 1994 as presented at the SCA Summer Conference on Collegiate Assessment (Litterst, VanRheenen, & Casmir, 1994). Not surprisingly, these surveys found that the deliberate teaching of oral communication skills is still neglected in U.S. public

schools. Therefore, the content domain of communication assessment probably cannot be defined by what is taught, but by what ought to be taught. The result would be that the specification of a content domain for testing would exert impact on instruction. That is, the instruments and administrative selection of an oral assessment tool or process can, in some cases, direct what may be included in the curriculum.

As such, the determination of what should be assessed is a pivotal decision. That decision can be informed by determining whether assessment of oral communication should take place with regard to the pedagogical process, impact on students, or any other of a variety of outcomes. Which of these three factors is, or are, of interest to the assessor? Is the process of teaching itself what should be assessed? Are we concerned with how we are doing what we are doing? Or is impact on students important to assess? Are we concerned with impact as measured by students' perceptions of a course or perhaps by students' improvement between pre- and posttest scores on a given communication assessment tool? Finally, is some other outcome factor such as retention of students, improved grade point averages, or even job placement of importance to the assessor?

Are the measures you choose to use or to develop valid?

When selecting an assessment procedure, one critical factor is validity. Does the procedure measure what it claims to measure? Does it relate to and measure purposes and objectives? With the possible exception of the Communication Competency Assessment Instrument (R. B. Rubin, 1982), most instruments for testing communication competence have not been subjected to studies of concurrent or predictive validity. In the time period since the publication of the first edition of this book, little work has been done in this area, especially in public education. One reason for this may be the frustration that state and local education agencies have experienced in attempting to set up assessment programs. The continued lack of accepted criteria against which tests may be validated constitutes another reason. Criteria for studies of predictive validity could include teacher or job ratings at some later point in time.

Establishing criterion-referenced validity seems particularly crucial in assessment tasks that are obviously contrived solely for the purpose of evaluation. Several assessment procedures require students to communicate in role-playing situations. While such procedures permit evaluation of "life role" communication skills, the relationship between role-playing performance and natural communication performance is still relatively unknown. R. B. Rubin (1985) and Morreale et al. (1993) have research data that support the relationship between the testing procedure and natural performance, but more research is needed.

Future research and development priorities relating to validity in measures of speaking and listening might include:

◊ establishing criterion measures for measuring concurrent and predictive validity,

◊ exploring naturalistic criterion measures for these purposes, or

◊ investigating criterion-referenced validity of contrived communication tasks.

Are assessment instruments reliable?

The field of speech communication has increasingly pursued investigations of test reliability since the first edition of this volume was published. As noted in the first edition, some attention had been given to internal consistency or dimensionality in studies of speech rating scales, and it appears to be common practice for commercial tests of listening skill to report this aspect of reliability. But other types of reliability need to be addressed more extensively. For example, measures of speaking proficiency (e.g., R. B. Rubin, 1982; Morreale et al., 1993) offer students a choice of topics, with no apparent evidence of equivalence between topics. Also, interviewer-assessor idiosyncrasies can alter performance in speaking assessments. Some writers have commented that single samples of speech are not reliable indicators of communication competence, and that several samples ranging over a variety of speech functions and situations should be taken for a fair assessment. At present, though, we lack the sort of precise information concerning the requisite size for a reliable sample. And insufficient information is available concerning test-retest reliability of most speaking and listening assessment instruments. Students do have good days and bad days. Instruction and practice may (ideally) change their skill over time. Both situations may weaken test-retest reliability results. Although individual test developers have done considerable work in establishing training procedures to support interrater

consistency and reliability, these procedures need to be instituted whenever and wherever the assessment of oral communication occurs. Guidelines for choosing appropriate statistical methods to evaluate interrater reliability are included in other SCA publications.

Future research and development priorities relating to reliability in measures of speaking and listening might include:

◊ determining the impact of varying topics and communication tasks on the testing process,

◊ determining the impact of assessor or test administrator on student performances,

◊ determining the size and diversity of speech samples required for a reliable indication of competence,

◊ ascertaining test-retest reliabilities of existing instruments, or

◊ supporting the institutionalization of methods to enhance interrater reliability.

What Measurement Techniques Are Presently Available?

Clearly the authors of this volume wish to encourage the development of new measurement techniques or the adaptation of new research methodologies for purposes of assessing oral communication. However, it is also worthwhile to examine some of the instruments already available. The instruments reviewed in Parts II and III are worth considering for use. In general, these instruments range on a continuum from direct (obtrusive) to indirect (unobtrusive) measures. Overall, the most optimal assessment procedures are those that are least intrusive. Naturalistic observation, and even classroom teacher ratings, however, introduce problems of rater bias and problems of consistency.

Indirect tests of speaking ability would alleviate many sources of inconsistency as they assess communication behavior as it naturally occurs within the classroom or context. However, it may be difficult to construct such tests. Moreover, indirect tests may be contaminated by extraneous factors like reading ability and "test-wiseness." Also, such indirect tests are likely to exert deleterious "washback" effects on speech communication instructional practices, and this can lead to a focus on rote knowledge rather than internalized skill.

Direct tests of listening ability present fewer problems involving consistency, particularly when test stimuli are tape-recorded. Even so, measures of listening in conversation are elusive. Interactive listening—where the listener is an equal conversational partner who responds and is ever ready to switch into the role of speaker—probably calls on different skills than procedures in which test-takers listen to a tape-recorded conversation, and calls on still other skills from procedures in which test-takers demonstrate their understanding of oral commands.

In tests of speaking ability, use of rating scales predominate. Typically, rating scales are applied to either extended talks or interview situations. It would be useful to adapt the use of rating scales to less intrusive, more interactive, communication situations.

Future research and development priorities relating to presently available measurement techniques might include:

◊ enhancing reliability of naturalistic observation procedures,

◊ developing more measures of communication in interactive situations, or

◊ extending performance rating scales to more interactive communication.

Are Assessment Instruments Inherently Susceptible to Group and other Biases?

Consistent group differences in test scores are not, in and of themselves, evidence of test bias. Rather, test bias can only be ascertained by determining if an instrument over- or under-predicts a particular group's performance on some criterion measure. Nevertheless, culture-bound evaluation materials will likely favor one cultural group over another. Such materials may include culture-bound communication contexts (e.g., role-playing a business executive), evaluation criteria (e.g., standard English pronunciation and syntax), and test stimuli (e.g., "Point to the grandfather clock.").

A less obvious source of potential bias against particular cultural groups is the very notion of oral communication assessment. For example, African-American children may construe the requirements of direct questioning by adults differently than do Anglo-American children. Similarly, socialization patterns among Native American Indians render an oral communication assessment as an anomalous

communication contest. In addition to biases against particular cultural groups, it is possible that communication assessment procedures may treat particular individuals differentially. Certainly individuals with organic speech defects should not be subjected to the same testing procedures as others. Additionally, certain personality traits may effect the communication assessment process. Most well-known among these is communication apprehension. Will special provisions be made for communication apprehension, or, if not, will the public schools be committed to "remediating" this condition as a part of their responsibility to prepare students for communication competency assessment?

Future research and development priorities relating to test bias in procedures for evaluating speaking and listening might include:

◊ developing criterion measures with which test bias may be determined;

◊ identifying culture-bound communication contexts, evaluation criteria, and stimulus materials;

◊ determining the degree to which oral communication assessment is inherently biased against particular cultural groups; or

◊ clarifying the status of personality traits vis-à-vis test bias.

Should Communication Competence Be Assessed?—A Postscript

One reason to assess communication competence, in addition to certifying proficiency among individual students, is to evaluate oral communication instruction. Of course, if our supposition of negligible communication instruction at some schools is accurate, then this motive is obviated. Another reason for assessment, therefore, is to encourage and guide the development of speech communication instruction. Testing tends to legitimize a teaching field, and test specifications may "washback" and stimulate instructional practice. Thus, important questions pertain to effects of communication assessment on teachers' and administrators' attitudes toward the legitimacy of speech communication, effects on curricular innovation, and effects on classroom practices.

One reason not to assess, and a potentially negative effect of any testing program, is the deterioration of student attitudes. Partly this is a function of the ends to which test results are put. If the results are used in beneficial ways for students as well as for teachers, then oral competency assessment is in order. In any event, it is worthwhile investigating whether the potential benefits of evaluating communication skills are offset by negative attitudinal outcomes.

Because large scale assessment of oral competency is just becoming widespread, limited information is available concerning its effects on institutional allocations. What are the costs of oral communication testing in terms of instructional hours, personnel hours expended, and dollars spent? It is indeed likely that many administrators do not encourage large scale direct measurement of speech communication competency because they fear it will be too costly. Such cost estimates must be weighed against presently unquantifiable benefits and utility to students, teachers, and administrators. SCA has clearly stated, through the delineation of assessment instrument criteria, that the students must either speak or listen (SCA, 1993). While the time and effort necessary to conduct effective speaking and listening assessment may seem daunting, the effort is well worth the benefits and outcome.

In summary, we propose the following research and development priorities relating to the utility and advisability of procedures for assessing oral communication competencies and speaking and listening:

◊ ascertaining whether measures are sensitive for purposes of assessing instructional impact;

◊ determining effects and value of assessment programs for students, teachers, and administrators;

◊ determining the curricular and instructional results of assessment programs;

◊ identifying the ends toward which test results are put; and

◊ ascertaining effects of evaluating students' competencies on their attitudes toward communication.

Conclusion

In spite of considerable advances made since the first edition of this volume, much work in the area of oral competency assessment remains to be done. Parts II and III review some available instruments and, in doing so, point out what is available and what is missing. Some gaps exist in gen-

eral measures of communication competency, as well as in other communication contexts. Speech communication professionals are presently at work filling these gaps by engaging in research on test development, validity, and reliability (Morreale, et al., 1994). The results of these efforts will be of great benefit to students and to the communication discipline. As professionals committed to the importance of oral communication and to the effective education of our nation's students, we have an opportunity through assessment to have a major impact on the priorities we value.

References

Allen, R. R., & Brown, K. L., Eds. (1976). *Developing communication competence in children*. Skokie, IL: National Textbook.

Backlund, P. (1985). Essential speaking and listening skills for elementary schools. *Communication Education, 34*, 186-195.

Backlund, P., Booth, J., Moore, M., Parks, A., & VanRheenen, D. (1982). A national survey of state practices in speaking and listening assessment. *Communication Education, 31*, 126-129.

Barnes, D. (1980). Situated speech strategies: Aspects of the monitoring of oracy. *Educational Research, 32*, 123-131.

Bassett, R. E., Whittington, N., & Staton-Spicer, A. (1978). The basics in speaking and listening for high school graduates: What should be assessed? *Communication Education, 27*, 293-303.

Berko, R. (1994). How will Goals 2000 affect speech communication? *SPECTRA, 30*(5), 1-11.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement, 5*, 115-124.

Curtis, D., Winsor, J., & Stephens, R. (1989). National preferences in business and communication education. *Communication Education, 38*, 6-14.

Dickson, W. P , & Patterson, J. H. (1979). Evaluating referential communication games for teaching speaking and listening skills. *Communication Education, 30*, 11-21.

Galvin, K., & Brommel, B. (1991). *Family communication: Cohesion and change* (3rd ed.). New York: Harper Collins.

Hay, E. A. (1992). *Program Assessment in oral communication*. Annandale, VA: Speech Communication Association.

Hurt, H. T., & Preiss, R. (1978). Silence isn't necessarily golden: Communication apprehension, desired social choice, and academic success among middle-school students. *Human Communication Research, 4*, 315-328.

Larson, C., Backlund, P., Redmond, M., & Barbour, A. (1978). *Assessing functional communication*. Annandale, VA: Speech Communication Association.

Litterst, J., VanRheenen, D., & Casmir, M. (1994). Practices in statewide oral communication assessment. In S. Morreale, R. Berko, M. Brooks, & C. Cooke (Eds.), *1994 SCA summer conference proceedings and prepared remarks* (pp. 187-215). Annandale, VA: Speech Communication Association.

Lundsteen, S. W. (1979). *Listening: Its impact on reading and the language arts*. Urbana, IL: ERIC and National Council of Teachers of English.

Maslow, A. H. (1954). *Motivation and personality*. New York: Harper & Row.

McCroskey, J. C., & Daly, J. A. (1976). Teachers' expectations of the communication apprehensive child in the elementary school. *Human Communication Research, 3*, 67-72.

Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.

Mead, N. A. (1982). *Assessment of listening and speaking performance*. Paper presented at the National Symposium on Education Research, Washington, DC.

Morreale, S., Berko, R., Brooks, M., & Cooke, C., (Eds.). (1994). *1994 SCA summer conference proceedings and prepared remarks*. Annandale, VA: Speech Communication Association.

Morreale, S., Moore, M., Taylor, P., Surges-Tatum, D., & Hulbert-Johnson, R. (1993). *The competent speaker*. Annandale, VA: Speech Communication Association.

National Education Goals Panel. (1992). *The national education goals report: Building a nation of learners*. Washington, DC. Author.

Piche, G. L., Rubin, D. L., & Turner, L. J. (1980). Training for referential communication accuracy in writing. *Research in the Teaching of English, 14*, 309-318.

Rubin, D. L. (1981). Using performance rating scales in large scale assessments of speaking proficiency. In R. Stiggins (Ed.) *Perspective on Oral Communication Assessment in the 80's*. Portland, OR: Northwest Regional Educational Laboratory.

Rubin, R. B. (1982). Assessing speaking and listening skills at the college level: The Communication Competency Assessment Instrument. *Communication Education, 31*, 19-32.

Rubin, R. B. (1985). The validity of the Communication Competency Assessment Instrument. *Communication Monographs, 52*, 173-185.

Rubin, R. B., & Graham, E. E. (1988). Communication correlates of college success: An exploratory investigation. *Communication Education. 37*, 14-27.

Rubin, R. B., Graham, E. E., & Mignerey, J. T. (1990). A longitudinal study of college students' communication competence. *Communication Education, 39*, 1-13.

Rubin, R. B., Moore, M. R., Sisco, J. I., & Quianthy, R. (1983). *Oral communication assessment procedures and instrument development in higher education*. Annandale, VA: Speech Communication Association.

SCANS Report for America 2000 (1993). *What work requires of schools*. Washington, DC: U.S. Department of Labor.

Speech Communication Association (1979a). Criteria for evaluating instruments and procedures for assessing speaking and listening. *SPECTRA, 15*, 5.

Speech Communication Association (1979b). *Standards for effective oral communication programs*. Annandale, VA: Author.

Speech Communication Association (1985). *Resources for assessment in communication*. Annandale, VA: Author.

Speech Communication Association (1990). *Communication is life: Essential college sophomore speaking and listening competencies*. Annandale, VA: Author.

Speech Communication Association (1993). The Speech Communication Association's criteria for the assessment of oral communication. Annandale, VA: Author.

Spitzberg, B. H., & Cupach, W. R. (1989). Handbook of interpersonal competence research. New York: Springer-Verlag.

Spitzberg, B. H. (1995). The conversational skills rating scale. Annandale, VA: Speech Communication Association.

Stiggins, J., Backlund, P. M., & Bridgeford, N. (1985). Avoiding bias in the assessment of communication skills. Communication Education, 34, 136-141.

Vangelisti, A., & Daly, J. (1989). Correlates of speaking skills in the United States: A national assessment. Communication Education, 38, 132-143.

Oral Communication Assessment Instruments for K-12

Neil J. O'Leary, *Dos Gatos Enterprises*
Ray J. Fenton, *Anchorage School District*

The assessment of communication skills has always played an important part in K-12* education. Teachers want students to improve their social skills, to consider and appreciate the feelings of others, to cooperate, to resolve conflicts, and to be effective public speakers and leaders. Because of these concerns, they spend much of their time providing instruction in communication, although they seldom identify their lessons as teaching communication.

Assessment of student communication by K-12 teachers has largely been informal and diagnostic. Communication skills instruction has been integrated across the curriculum and has not received the recognition or formal evaluation that is much more common in the traditional "three Rs." Reading, writing, and arithmetic are prominent on the report card and are more subject to formal evaluation through standardized multiple-choice tests and performance assessments. Communication appears occasionally as a secondary course or as part of a language arts unit. There has, however, been an increased recognition by states and school districts that communication, and the assessment of communication, is important to success in both school and the workplace (SCA, 1994).

This introduction to reviews of formal K-12 assessment measures of communication briefly examines trends in the development of communication assessment. It then examines the increasing emphasis on performance assessment and the Speech Communication Association (SCA, 1993) standards for the assessment of oral communication. Finally, it outlines common sources which may be used in locating communication assessment measures, and the criteria that were used for the selection of the 16 instruments included in this compendium.

Trends: K-12 State Standards and Assessment Efforts

In recent years, more formal assessment of communication skills has been encouraged through the development of goals by many states and school districts. Twenty-three states now consider oral communication to be part of a comprehensive language arts program, and most have written goals or performance standards in oral communication. A recent Speech Communication Association survey of the status of oral communication as a state education requirement found that 25 states require and 15 states recommend oral communication as part of the K-12 program (Chesebro, 1990).

Most of the states and school districts with articulated standards have not gone on to develop formal measures of oral communication competencies. Instead, many have developed model curricula with informal measures or suggested assessment activities. Some states call for local education agencies either to develop courses or to conduct assessment that meets state guidelines. A few others have adopted or developed standardized measures of listening, specific performance assessment requirements in secondary communication classes, or specified certain formal performances.

The chart below provides a breakdown of the status of speech communication assessment processes by state. The information is adapted from Chesebro (1990), VanRheenen and Casmir (1990), and Fenton and O'Leary (1993). It includes states that are encouraging local assessment, states that have developed or are in the process of developing assessment processes, and states that have not developed assessment processes as yet.

No state assessment processes meet the standards set out by the Speech Communication Association for communication instruments (SCA, 1993). However, the substantial work done by state departments of education and state educators in developing standards, curriculum frameworks, model lessons, model assessment processes, and rating instruments makes them fertile ground for seeking exemplar instruments. If the refinement and improvement of state-developed communication assessment instruments continues, more reliable and valid assessment instruments should become widely available.

11* Grade levels are approximate. Some measures provide age rather than grade levels.

Status, by State, of Speech Communication Assessment Programs*

Assessment Encouraged	Georgia, Hawaii, Idaho, Illinois, Indiana, Iowa, Massachusetts, Michigan, Minnesota, Montana, New Mexico, Ohio, Oregon, Pennsylvania, Rhode Island, Texas, Utah, Vermont
Assessment Developed or Developing	Alaska, Arizona, California, Connecticut, Kansas, Kentucky, New York, North Carolina, South Dakota, Virginia, West Virginia
Assessment Not Developed	Alabama, Arkansas, Colorado, Delaware, Florida, Louisiana, Maine, Maryland, Mississippi, Missouri, Nebraska, Nevada, New Jersey, New Hampshire, North Dakota, Oklahoma, South Carolina, Tennessee, Washington, Wisconsin, Wyoming

* A more detailed description of state requirements and mandates for oral communication is provided elsewhere (SCA, 1994).

Performance Assessment: A Growing Emphasis at K-12

Many of the states that have developed model assessment programs are following the national movement away from traditional machine-scored, short-answer measures toward performance assessments based on observation and expert judgment. This movement is consistent with the long tradition of diagnostic assessment in the field of communication and is a reflection of increasing recognition that short-answer measures are poor measures of complex performances, such as those that are common in effective interpersonal communication.

Performance assessment requires students to respond to broader, more-complex problems in order to reflect authentic behavior in real situations. The rating of multiple characteristics of complex performances has been a hallmark of assessment in public speaking, interpersonal communication, and small group communication. Examples of some typical performance assessment methods used in comprehensive language arts programs are found in Bird, Goodman, and Goodman's (1994) *The Whole Language Catalogue: Forms for Authentic Assessment*. These assessments are reflected in the programs that are becoming more common in the local school districts (Matter & O'Reilly, 1993).

The difficulty in developing adequate performance assessment measures stems from the diversity and complexity of communication behaviors. The wide range of performance expectations in the communication classroom can be found in Barbara Wood's *Oral Communication in the Elementary Classroom* (Thais & Suhor, 1984) and the Speech Communication Association's (1991) *Guidelines for Developing Oral Communication Curricula in Kindergarten through Twelfth Grade*. Such complex expectations suggest that the portfolio approaches of the language arts and authentic assessment techniques are uniquely suited for assessment of oral communication instruction. Unfortunately, there has been very little development of authentic assessment strategies and portfolio assessments in communication (Arter, 1994).

Alternative assessment instruments, checklists, rating forms, and communication portfolios, which provide useful informal assessment, do not generally meet the high standards set out by the Speech Communication Association (1993) for assessment instruments. While many examples of valuable resources for informal assessment are found in textbooks and accompanying teacher guides (see, for example, Galvin, Cooper, & Gordon, 1994), these are not presently included in this review because they lack formal validation.

SCA Standards for Assessment

The Speech Communication Association (1993) has advanced a set of standards for the judgment of assessment instruments, that can be viewed as ideal criteria for developing an assessment program. In terms of relevancy for the K-12 grade levels, the SCA standards may be paraphrased as follows:

General standards

◊ Assessment should include measures of knowledge, assessment of skills, and evaluation of attitudes toward communication.

◊ Assessment should consider the judgments of trained assessors as well as the self-reflections of the individuals being assessed.

◊ Assessment should clearly distinguish speaking and listening skills from reading and writing skills.

◊ Assessment should be sensitive to the effects of relevant physical and psychological disabilities of the individual being assessed.

◊ Assessment should be based on atomistic/analytic data as well as on holistic impressions.

Assessment content

Assessment should consider both verbal and nonverbal aspects of communication and should consider competence in more than one communication setting.

Assessment should consider specialized fields appropriate to the course of study or the specialty of the person being assessed.

Assessment instruments

◊ Instruments should be consistent with the dimension of oral communication being addressed.

◊ Instruments should describe degrees of competence.

◊ Instruments should clearly identify the range of responses that constitute various degrees of competence. Examples of various degrees of competence should be included as anchors.

◊ Instruments should possess an acceptable level of reliability.

◊ Instruments should possess appropriate validity.

◊ Instruments should be free from cultural, sexual, ethical, racial, age, and developmental bias.

◊ Instruments should be suitable for the developmental level of the individual being assessed.

◊ Instruments should be standardized and detailed enough that individual responses will not be affected by the administrator's skill in administration.

Procedures and administration

◊ Assessment procedures should protect the rights of individuals.

◊ Assessment procedures, which result in classification of individuals, should be based on multiple sources of information.

◊ Assessments should only be conducted by individuals who have had sufficient training.

SCA (1993) standards further state that "periodic assessment of oral communication competency should occur annually during the educational careers of students," and "an effective systematic assessment minimally should occur at educational levels K, 4, 8, 12, 14, and 16" (p. 4).

There are no comprehensive assessment instruments for grades K-12 that meet all of the standards set by the Speech Communication Association. The standards related to demonstrated reliability and validity, standardization of administration, freedom from bias, and utility for the classification of individuals and programs to eliminate most of the available assessment measures from consideration. These high SCA standards set out the characteristics of ideal measures and may be used as criteria that may be the goal to strive toward when planning an assessment program.

Identifying Measures

Several reviews and lists of communication assessments are available. Principal among those that provide critiques of various widely-available measures are Rubin and Mead's (1984) *Large Scale Assessment of Oral Communication Skills: Kindergarten Through Grade 12*, Arter's (1989) *Assessing Communication Competence in Speaking and Listening*, and Educational Testing Service's (1991) *Measures of Speech/Communication*. Additional publications provide insights into various conceptual, methodological, and practical concerns that must be considered in the assessment of oral communication. Principal among these more general works are Thaiss and Suhor's (1984) *Speaking and Writing K-12* from the National Council of Teachers of English; Underhill's (1987) *Testing Spoken Language*, which examines performance assessment from a British perspective; and Spitzberg and Cupach (1988) *Handbook of Interpersonal Competence Research*, which provides an accumulation of measures from various disciplines. Taken together, these published sources provide an excellent overview of existing measures and the problems associated with the development of standardized assessment programs for human communication.

Knowledgeable language arts specialists work in almost every state and provincial department of education in North America. In the compilation of this introduction and its compendium of instru-

ments, all state and provincial education agencies were contacted. Many provided copies of standards, model curricula and assessment programs, and assessment instruments. Additionally, letters were sent to the presidents of all state and regional speech communication associations; some printed notices of the search were placed in their newsletters. A request for instruments was printed in the national Speech Communication Association newsletter, *Spectra*.

Searches were made of the ETS test catalogue (gopher.cau.edu); the ERIC system; and the National Center for Research on Evaluation, Standards, and Student Testing's Alternative Assessments in Practice Database (spinoza.cse. ucla.edu). Requests for information were made to the Test Collection at Northwest Regional Educational Laboratory, the ERIC Clearinghouse and Test Collection at Catholic University of America, and Buros Institute at the University of Nebraska- Lincoln. Letters were also sent to major test publishers and to the authors of instruments included in earlier summaries of communication assessment (Arter, 1989; Rubin & Mead, 1984).

More than 3,000 oral communication and listening measures were located and screened for potential inclusion in this compendium. The initial screening eliminated measures that are not readily available and that had no published indicators of validity or reliability. Only 98 instruments were selected for more careful examination. The 98 instruments were solicited from publishers and authors or were gleaned from publications. These were subjected to further screening based on the following set of criteria.

Selecting Measures

Instruments were selected for review based on a set of criteria that evolved from an examination of the SCA (1993) standards for assessment. In making the decisions regarding inclusion, the following criteria were applied:

◊ The instrument must, at least in part, directly measure performance. Reports of attitude, self-assessment, or self-reflection are included if part of the assessment package includes performance.

◊ Valid relationships must be demonstrated between the skill assessed and important communication constructs or K-12 instructional goals.

◊ Activities and administration must be standardized, and directions to administrators and students must be sufficient.

◊ Performance rating forms must have clear rubrics, benchmarks, or sample performances to assure reliable scoring. Where training of raters is required, adequate training materials must be available.

◊ The instrument must be useful for diagnosis and/or program assessment.

◊ Instruments must be useful for normal populations. Instruments developed for screening for special education students and for assessing speech and language or psychological disorders are excluded unless they may be useful with normal populations and administered by classroom teachers.

◊ The instrument and all materials needed for administration and scoring must be publicly available.

Application of these criteria eliminated many instruments that would have valid instructional uses in the K-12 classroom. The focus on skills and the exclusion of attitude surveys eliminated self-report instruments, such as commonly-used measures of communication apprehension, socialization, family behavior, self-disclosure, and reports of general communication or listening experience. Instruments that measure progress or ability of students with disabilities are excluded. Psychological measures are excluded, although some of them clearly focus on variables that affect interpersonal interaction. Finally, academic achievement tests (those that are orally presented) and measures of sight reading skills (those measures that are not direct assessments of communication or listening skills) are also not considered. Instruments without some clear indication of reliability and validity were also eliminated from the review. The need for formal indicators of reliability and validity come to the fore when assessment is used to assign grades, measure program success, compare programs, and measure program or individual performance against predefined standards.

Only 16 instruments satisfactorily met all of the criteria outlined above. None of the 16 fully meets the SCA (1993) guidelines, nor may they be considered part of a system that provides for the assessment of growth across grades K-12. Many of the instruments only measure one or two aspects of communication. None includes both assessment of performance by observers and self-assessment

by those who are the subjects of observation. However, the 16 measures included in the following compendium were found to fulfill most of the SCA (1993) standards for measures of oral communication.

The paucity of satisfactory measures calls attention to the need to continue to develop and evaluate tools for the assessment of oral competency for grades K-12. Future editions of this volume may include self-report and other measures that are clearly valuable as part of instruction, even though they may lack standardization and demonstrated validity and reliability.

Psychometric Properties

Each review includes an assessment of the validity, reliability, and potential ethnic/gender bias of instruments. In most cases, the discussion of psychometric properties is based on published information made available from the instrument developer, most often in a technical or administration manual. In a few cases, additional information is included from research articles or other sources, listed as references following the review.

Validity is the degree to which the measure actually measures the content, trait, or construct that it purports to measure. Reliability is the extent to which the measure yields consistent results. Bias is the extent to which extraneous factors such as race or gender result in differing assessment of performance for individuals who actually have the same ability.

Conclusion

Rubin and Mead commented in 1984 that "despite this initial impetus to evaluate communication competencies, despite this view that developing assessment procedures presents no insurmountable technical obstacles, and despite some concrete suggestions of pertinent measurement instruments, attempts to implement large scale assessment of speaking and listening skills have not been forthcoming" (p. 3). There has not been sufficient improvement in the assessment field in the ensuing period.

Despite some admirable developments since 1984, which have included a more detailed and complete definition of the K-12 communication curricula (SCA, 1991) and the development of standards for measures of oral communication (SCA, 1993), there continues to be a need for more comprehensive assessment instruments. Teachers and researchers are not given enough choice among instruments that have proven themselves as valid and reliable measures of the communication behaviors necessary for success in school and in life beyond school. There are no instruments that allow us to assess students with regard to the goals and standards articulated by state departments of education and school districts. And, there is a growing need as the emphasis on accountability, assessment of standards, and professional training of teachers of communication grows (Most, 1994).

It is the hope of the reviewers of instruments in this compendium that soon we will be able to revisit the question of the availability of K-12 measures and find better-developed assessment instruments and programs. The many informal instruments, which presently lack evidence of validity and reliability, can serve as a base for the development of more useful measures.

References

Arter, J. A. (1989). Assessing communication competence in speaking and listening: A consumer's guide. Portland: Test Collection, Northwest Regional Educational Laboratory.

Arter, J. A. (1994). [Personal communication, given to the Portland, Oregon, Northwest Regional Educational Laboratory, Test Collection, concerning sample materials developed in various Oregon school districts related to speech communication portfolios]. Unpublished raw data.

Bird, L. B., Goodman, K. S., & Goodman, Y. M. (1994). *The whole language catalogue: Forms for authentic assessment*. New York: Macmillan/McGraw-Hill School Publishing Company.

Chesebro, J. W. (1990). *Final data tables: The status of oral communication in the United St06/28/96ates as a state education requirement*. Annandale, VA: Speech Communication Association.

Educational Testing Service. (1991). *Measures of speech/ communication*. Princeton, NJ: Author.

Fenton, R., & O'Leary, N. (Eds.). (1993, February). *Report of the ad hoc committee on the status of oral communication requirements in the western United States*. Paper presented at the meeting of the Western Communication Association, Albuquerque.

Galvin, K. M., Cooper, P. J., & Gordon, J. M. (1994). *Teacher's resource book. The basics of speech: Learning to be a competent communicator*. Lincolnwood, IL: National Textbook Company.

Matter, K., & O'Reilly, J. (1993). *Proceedings of the 1993 annual symposia presented at the national council on measurement in education annual conference*. Portland, OR: National Association of Test Directors.

Most, M. G. (1994). Certification standards for speech communication teachers: A national survey. *Communication Education, 43*, 195-204.

Rubin, D. L., & Mead, N. A. (1984). *Large scale assessment of oral communication skills: Kindergarten through grade 12*. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills.

Speech Communication Association. (1991). *Guidelines for developing oral communication curricula in kindergarten through twelfth grade*. Annandale, VA: Author.

Speech Communication Association. (1993). *Criteria for the assessment of oral communication*. Annandale, VA: Author.

Speech Communication Association (1994). *SCA summer conference on collegiate oral communication competency: Published proceedings*. Annandale, VA; Author.

Spitzberg, B., & Cupach, W. R. (1988). *Handbook of interpersonal competence research*. New York: Springer-Verlag.

BEST COPY AVAILABLE

Thaiss, C., & Suhor, C. (Eds.). (1984). *Speaking and writing K-12*. Urbana, IL: National Council of Teachers of English.

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Cambridge, England: Cambridge University Press.

VanRheenen, D., & Casmir, M. (1990). *Current practices in communication assessment*. Los Angeles: Pepperdine Univ.

Selected Bibliography

Alberta Education. (1990). *Oral communication evaluation: English 30/33-activities and scoring guides*. Edmonton, Alberta: Student Evaluation Branch, Alberta Education.

Cooper, C. R. (1981). *The nature and measurement of competency in English*. Urbana IL: National Council of Teachers of English.

Educational Testing Service. (1993). *The ETS test collection catalog, 2*. Phoenix: Oryx Press.

Fagon, W. T., Cooper, C. R., & Jensen, J. M. (1985). *Measures for research and evaluation in the English language arts, 2*. Urbana, IL: National Council of Teachers of English.

Gray, P. A. (1989). Improving speech communication assessment K-12. In P. J. Cooper & K. M. Galvin (Eds.), *The future of speech communication education* (p. 96). Annandale, VA: Speech Communication Association.

Illinois State Board of Education. (1982). *Assessing oral communication skills*. Springfield, IL: Author. (ERIC Document Reproduction Service No. ED 222 937).

Larson, C., Backlund, P. M., Redmond, M., & Barbour, A. (1978). *Assessing functional communication*. Annandale, VA: Speech Communication Association.

Litterst, J. K. (1992). How the state association can make a difference in communication assessment. *Communication and Theater Association of Minnesota Journal, 19*, 1-9.

Massachusetts State Department of Education. (1981). *Speaking Assessment. Final Report*. Boston: Bureau of Research and Assessment. (ERIC Document Reproduction Service No. ED 210 238).

Mitchell, J. V. (Ed.). (1992). *The eleventh mental measurement yearbook*. Lincoln: University of Nebraska.

New York State Education Department. (1975). *Listening and speaking, K-3: A packet for teachers*. Albany, NY: Bureau of Elementary Curriculum Development. (ERIC Document Reproduction Service No. ED 132 594).

Ohio Department of Education. (1992). *Competency-based language arts program*. Columbus, OH: Division of Curriculum, Instruction, and Professional Development, Ohio Dept of Education.

Oregon Department of Education. (1988a). *English language arts: Assessing student progress on the common curriculum goals, report 2, speaking skills*. Salem: Division of Curriculum and School Improvement, Oregon Dept of Education.

Oregon Department of Education. (1988b). *English language arts: Assessing student progress on the common curriculum goals, report 3: Listening skills*. Salem: Division of Curriculum and School Improvement, Oregon Dept of Education.

Rubin, D. L., Daly, J., McCroskey, J. C., & Mead, N. A. (1982). A review and critique of procedures for assessing speaking and listening skills among preschool through grade twelve students. *Communication Education, 31*, 285-303.

Snively, M. (1989). To what extent can there be objectivity in measuring speech-theatre competency? In P. J. Cooper & K. M. Galvin (Eds.). *The future of speech communication education* (pp. 72-73). Annandale, VA: Speech Communication Association.

Stiggins, R. (Ed.). (1981). *Perspectives on the assessment of speaking and listening for the 1980's*. Portland: Northwest Regional Educational Laboratory.

Tchudi, S. (1991). *Planning and assessing the curriculum in English language arts*. Alexandria, VA: Association for Supervision and Curriculum Development.

Reviewers' Note

Two good sources of K-12 instruments for review and information are the ERIC Clearinghouse on Assessment and Evaluation, Catholic Univ of America, 209 O'Boyle Hall, Washington DC 20064, 202-319-5120; and The Test Center, NW Regional Educational Laboratory, 101 SW Main #500, Portland OR 97240, 503-275-9500.

Editors' Note

In the following reviews of instruments for K-12 grade levels, the authors strove for consistency and uniformity in the presentation of each measure/assessment tool. However, variability exists across the instruments in terms of what information is and is not accessible and available. For example, not every instrument has been subjected to exactly the same testing procedures, so the same testing information is not available for each measure. In some cases, cost or other information is only available to the user by contacting the developer regarding his or her precise user needs.

Considering this variability, the reviewers chose to include information that was available, and to indicate "not available" after any item to which they had no access. If users of this volume have information that could enrich the presentation of any assessment tool, it should be sent to the authors, Neil J. O'Leary and Ray J. Fenton.

Oral Communication Assessment Instruments: K-12*

Title	Source	Context	Population	Administration	Page
Listening Instruments					
Brown-Carlson Listening Test	Harcourt, Brace, & World (1955)	listening	HS, adult	multiple-choice, paper-pencil	28
Kentucky Comprehensive Listening Test	Robert N. Bostrom (1990)	listening	HS, college	audiotape, respond on multiple-choice sheet	29
Listening Assessment for ITBS[]	Riverside Publishing (1992)	listening	grades 3-9	multiple-choice, paper-pencil	30
Listening Assessment for TAP/ITED	Riverside Publishing. (1992)	listening	grades 9-12	multiple-choice, paper-pencil	31
Metropolitan Achievement Test—7th Ed.	Psychological Corp. & Harcourt Brace Jovanovich (1992)	listening	grades K-1.5	multiple-choice, paper-pencil	32
Michigan Educational Assessment Program (MEAP) Listening Test	Michigan Dept of Education (1986)	listening	grades 4, 7, 10	audiotape, respond on multiple-choice sheet	33
National Achievement Test—2nd Ed.	American College of Testing (1990)	listening	grades K-2	items are read, subjects respond in booklet	34
Stanford Achievement Test—8th Ed.	Psychological Corporation (1991)	listening	grades K-9	sentences are read, subjects respond in booklet	35
Survey of Early Childhood Abilities (Test of Auditory Perceptual Skills	Psychological & Educational Publications (1987)	listening	ages 4-7	items are read, subjects respond on a test profile	36
Watson Barker Listening Test—HS Version	Spectra, Inc. (1989)	listening	HS	audio- or video- tape, respond on multiple- choice test	37
Interpersonal Instruments					
Diagnostic Analysis of Nonverbal Accuracy Scale	Stephen Nowicki (1992)	interpersonal	grades 1-5	group & individual admin., using slides	38
Test of Pragmatic Skills— Rev. Ed.	Communication Skill Builders (1986)	interpersonal	ages 3-8	individual admin. using puppets, blocks, telephones	39
Public Speaking Instruments					
Oral Proficiency Examination	Gary Community Schools Corp. (1977)	public speaking	grade 10	subjects interviewed by 3-person committee; interview is taped and scored by testers	40
Various Dimensions of Communication					
Communication Competency Assessment Instrument—HS	Spectra, Inc. (1993)	public speaking, interpersonal	grades 9-12	subjects present talk & engage in guided conversation with tester who scores responses	41
Hunter-Grundin Literacy Profiles	The Test Agency	spoken language	ages 6.5-12.7	subjects describe lunchroom scene, are rated on 5 language scales	42
Joliet 3-Minute Speech & Language Screen— Rev. Ed.	Communication Skill Builders (1993)	communication competency	grades K, 2, 5	individual administration using illustrations of objects	43

* These tools assess skills/behaviors and can be used for program and/or outcome assessment purposes.

BEST COPY AVAILABLE

Brown-Carlson Listening Test

Authors/developers: J. I. Brown, & G. R. Carlson. Sample for review available from Northwest Regional Educational Laboratory, 101 SW Main St #500, Portland OR 97204.

Source: Harcourt, Brace, & World (1955). [Although this test is no longer in print, it is still widely available.]

Cost: Information not available.

Intended context or behavior: Listening.

Intended academic level: High School/College/Adult.

Time required: 50 minutes.

Test description: The Brown-Carlson Listening Test involves 76 multiple-choices items that are read to students by the teacher. The test includes sections about words, short passages, and paragraphs. Students respond on separate answer sheets. Skills assessed include short-term recall, following directions, recognizing transitions, recognizing word meaning, and lecture comprehension.

Scoring: This is a multiple-choice test that is hand-scored.

Norm/criterion reference populations: Original norming population included 8,000 high-school students who were representative of the national population in ability and age.

Culture or gender biases: Items do not appear to have a gender/ethnic bias. No information is presented on the interaction between race/dialect/ethnicity of the presenter of the test and test takers. The test has been used with a wide variety of groups.

Validity

Predictive: Moderate correlations with high-school rank (.21, .28) and with honor-grade points (.48).

Concurrent: Strong correlations with tests of mental ability, above .70. Correlations with reading tests ranged from about .50 to about .70.

Face/content: Items were selected based on criteria generated by the authors and experience with field tests. Word meaning and lecture comprehension items appear to be related to activities that are precursors to success in school.

Reliability

Alternate forms: Forms Am and Bm correlate .78.

Internal consistency: Split-half correlations reported of .84 to .90.

Evaluative Reactions

Practicality and usefulness: Widely used measure, used frequently in listening research. The test is easy to administer.

Clarity of instructions: Instructions are clear.

Range of responses or degree of competence: Scales provide adequate differentiation of levels of performance.

Primary strengths: Ease of administration and wide use.

Primary weaknesses: Because the test is out of print, it is increasingly difficult to obtain. Comparison groups for norming are dated.

Overall adequacy: The division of results into scales related to various skills makes the test a useful instructional aid.

Additional References

Bateman, D., Frandsen, K., & Dedmon, D. (1964). Dimensions of lecture comprehension: A factor analysis of listening test items. *Journal of Communication, 14*, 183-189.

Brown, J. I., & Carlson, G. R. (1955). *Brown-Carlson Listening Comprehension Test*. New York: Harcourt Brace and World.

Kelley, C. M. (1965). Investigation of the construct validity of two commercially published listening tests. *Speech Monographs, 32*, 139-143.

Reviewer: Ray J. Fenton.

Kentucky Comprehensive Listening Test

Author/developer: R. N. Bostrom. The Kentucky Listening Research Center, PO Box 555, Lexington KY 40506. 606-257-7800.

Source: Bostrom, R. (1990). *Listening behavior: Measurement and applications*. New York: Guilford.

Cost: Sample Packet, including 30 answer sheets and tape for Form A, scoring key and descriptive booklet, $60; Complete kit, with materials for Form A and Form B, $150.

Intended context or behavior: Listening.

Intended academic level: University and high-school students—grades and levels are not differentiated.

Time required: About 30 minutes.

Test description: The test consists of four segments testing the ability of individuals to (a) retain information for very brief periods; (b) retain material in a "rehearsal" mode for short periods; (c) draw conclusions based on the emotional tone, context, and content of a dyadic interaction; and (d) retain information presented from a lecture. Stimuli items are presented on an audiotape, and listeners select responses on a multiple-choice answer sheet. Some items include background conversation as a distraction.

Scoring: There are scores for Short Term Listening, Short Term Listening/Rehearsal, Interpersonal Listening, Lecture, Distraction, and Total. Multiple-choice forms can be machine- or hand-scored.

Norm/criterion reference populations: Percentile rank scores are based on a university student population. The exact nature of the norming population is not indicated in the accompanying booklet. Additional information is available in the source article. More than 3,700 students were included in the most recent norming population which is described as "a good cross-section of the undergraduate population." Materials indicate that the performance of high school and college students is similar.

Culture or gender biases: There is no explicit discussion of gender or ethnic bias. Some research is reviewed suggesting small differences in male and female listening behavior.

Validity

Concurrent: Moderate to low correlations are reported with ACT English (.41), math (.39), social studies (.40), natural sciences (.23).

Face/content: Short-term memory of series of numbers and letters may or may not have a connection with effective interpersonal communication. Assessing the emotional state of speakers and the underlying meaning of emotional comments are clearly related to interpersonal and educational communication activities. Recalling information from a lecture is also clearly related to communication in an educational environment.

Reliability

Alternate forms: Two forms of the test are available inadequate information is included to judge the test-retest reliability of the forms.

Test/retest: Test-retest reliability on an earlier version of the test ranged from .78 to .87 on the Short Term Listening, Short Term Listening—Rehearsal, and Lecture scales.

Internal consistency: Cronbach alphas are reported for a number of test administrations as above .63 for Short Term Listening, above .53 for Short Term Listening with Retention, above .46 on Interpersonal; above .71 on Lecture; above .46 with Distraction; and from .71 to .83 overall.

Evaluative Reactions

Practicality and usefulness: The instrument is easily administered and appears to have reasonable reliability for use in communication research. The authors provide information and a Listening Profile, which would prove useful to students attempting to understand their own listening ability and to teachers who wish to make an informal assessment of listening as part of developing instructional activities.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Directions are clear and call for a check on the ability of students to hear and understand the speaker on the tape.

Primary strengths: An easily administered test with standardized prompts that allows assessment of various types of listening related to the interpretation of conversation and retention of information from a lecture. Prior use provides a good indication that the test has works with various groups of students and adults.

Primary weaknesses: The relationship between Short Term Listening and Short Term Listening/Rehearsal to effective interpersonal communication and retention of information such as that presented in lectures and classroom discussions is not clear.

Overall adequacy: An excellent instrument as part of instruction that includes the examination of listening skills in various contexts.

Additional References

Rubin, R. B., & Roberts, C. V. (1987). Comparative examination and analysis of three listening tests. *Communication Education, 36*(2), 142-153.

Rubin, R. B., & Shepherd, P. (1985). *Listening assessment via the Communication Competency Assessment Instrument*. Kent, OH: Kent State University. (ERIC Document Reproduction Service No. ED 264 616)

Rubin, R. B., & Shepherd, P. (1985, November). *Listening assessment via the Communication Competency Assessment Instrument*. Paper presented at the annual meeting of the Speech Communication Association, Denver, CO.

Reviewer: Ray J. Fenton.

Listening Assessment for ITBS

Authors/developers: A. N. Hieronymus, H. D. Hoover, K. R. Oberly, & N. Cantor. Riverside Publishing Co., 8420 Bryn Mawr Ave, Chicago IL 60631.

Source: Riverside Publishing Company (1992).

Cost: Classroom set of 50, with materials for machine scoring, $42.40; Interpretive Guide for Listening, $4.50; scoring is extra. Test may be hand-scored.

Intended context or behavior: Listening.

Intended academic level: Grades 3-9.

Time required: 35 minutes.

Test description: The test is made up of 95 questions with between 31 and 40 questions administered for each level. Levels 9-14 overlap and include questions and passages of increasing difficulty, which are presented by the teacher. Objectives assessed at each level include: Literal Meaning, Inferential Meaning, Following Directions, Linguistic Relationships, Numerical/Spatial/Temporal Relationships, and Speaker's Point of View or Style.

Scoring: Multiple-choice answer sheets may be machine- or hand-scored.

Norm/criterion reference populations: Normed with the Iowa Tests of Basic Skills. Grade equivalents, standard scores, and percentile ranks are available. School average norms are available for group comparisons.

Culture or gender biases: Iowa test items are reviewed by a panel for potential bias and subjected to statistical analysis which identifies questionable items. Items are carefully screened for any language or content bias.

Validity

Face/content: Items on the various levels of the test relate to common classroom and real world listening activities. Presentation of items by the teacher is positive because the speaker is familiar, but this may be a limiting factor in generalization of scores across situations.

Construct: Items directly relate to the objectives identified by the authors.

Reliability

Internal consistency: KR-20 reliability coefficients range from .70 to .81 for the various levels of the listening test. While this is lower than coefficients for other ITBS subtests, it is good for a short test where multiple, complex skills are tested.

Evaluative Reactions

Practicality and usefulness: Easy to administer listening test that reflects a broad range of listening skills. Unusually complete reports which may be included with reports ITBS achievement test scores.

Clarity of instructions: Instructions are clear and easy to follow.

Standardization in administration and evaluation: Multiple-choice responses. Differences in administration may affect student scores. The familiarity with the teacher who administers the test may result in overestimation of the ability of student when listening to unfamiliar speakers.

Primary strengths: Ease of administration, range of objectives, and quality of questions and materials.

Primary weaknesses: Test does not include conversational situations.

Overall adequacy: A high-quality standardized listening test that should be given consideration. Valuable when used in conjunction with the ITBS skills tests as a means of gaining insight into student performance. Listening subtests are included in the complete battery of Iowa Tests of Basic Skills for grades K-3.

Additional References

Deeter, T., & Wilson, M. (1993). *Assessment program results 1992-1993*. Des Moines: Des Moines Public Schools. (ERIC Document Reproduction Service No. ED 370 972)

Han, L., & Hoover, H. D. (1994). *Gender differences in achievement test scores*. Iowa City: University of Iowa. (ERIC Document Reproduction Service No. ED 369 816)

Han, L., & Hoover, H. D. (1994, April). *Gender differences in achievement test scores*. Paper presented at the meeting of the National Council of Measurement in Education, New Orleans.

Magliacano, K. (1994a). *The effect of a child's age at school entrance on reading readiness and achievement test scores*. Bloomfield, NJ: Kean College. (ERIC Document Reproduction Service No. ED 366 939)

Magliacano, K. (1994b). *The effect of a child's age at school entrance on reading readiness and achievement test scores*. Unpublished master's thesis, Kean College, Bloomfield, NJ.

Reviewer: Ray J. Fenton.

Listening Assessment for TAP/ITED

Authors/developers: O. M. Haugh, & D. P. Scannell. Riverside Publishing Co., 8420 Bryn Mawr Ave, Chicago IL 60631.

Source: Riverside Publishing Company (1992).

Cost: Classroom set of 50, with materials for machine scoring, $36; Interpretive Guide for Listening, $4.50; scoring is extra. Test may be hand-scored.

Intended context or behavior: Listening.

Intended academic level: Grades 9-12.

Time required: 40 minutes.

Test description: The total test contains 100 questions with between 48 and 52 questions administered by the teacher for each level. Levels 15-18 overlap and include questions and passages of increasing difficulty. Objectives assessed at each level include: Literal Meaning (short-term memory of details about persons, objects, time, and place); Inferential Meaning (word meaning in context, sequential relations, and drawing conclusions); Speakers Point of View and Purpose (main point of lecture); Differences Between Fact and Opinion; and Detecting Bias.

Scoring: The multiple-choice answers may be machine- or hand-scored.

Norm/criterion reference populations: Normed with the Iowa Tests of Educational Development and Tests of Achievement and Proficiency. Grade equivalents, standard scores, and percentile ranks are available. Group average norms are available.

Culture or gender biases: All items were reviewed by a panel for content or language that would affect student performance. Statistical analysis was done to identify items with differential performance for various groups.

Validity

Face/content: Items on the various levels of the test relate to common listening activities (e.g., lectures). Presentation of items by the teacher is positive because the speaker is familiar, but this may be a limiting factor in generalization of scores across situations.

Construct: The items directly relate to the listening skills used in school and are categorized according to four levels of cognition: knowledge/information, comprehension, application/analysis, and synthesis/evaluation. Criterion-referenced information appears to be directly related to listening and comprehension.

Reliability

Internal consistency: KR-20 reliability coefficients ranged from .83 to .85 for the various levels in a fall administration and from .82 to .85 in a spring administration.

Evaluative Reactions

Practicality and usefulness: The test is easily administered and directly relates to listening in an academic environment.

Clarity of instructions: Instructions are clear and well-written.

Standardization in administration and evaluation: Multiple-choice test. Administration is by the teacher and familiarity with the administrator may affect scores.

Primary strengths: Easily administered norm-referenced measure that provides individual and group averages.

Primary weaknesses: The test explores a limited situation without attention to listening to ongoing conversations or discussions.

Overall adequacy: Excellent example of a well-made norm-referenced test. Useful for assessing listening in an educational setting. Scores may be related to achievement scores on companion TAP or ITED instruments.

Additional References

Deeter, T., & Wilson, M. (1993). *Assessment program results 1992-1993*. Des Moines: Des Moines Public Schools. (ERIC Document Reproduction Service No. ED 370 972)

Han, L., & Hoover, H. D. (1994). *Gender differences in achievement test scores*. Iowa City: University of Iowa. (ERIC Document Reproduction Service No. ED 369 816)

Han, L., & Hoover, H. D. (1994, April). *Gender differences in achievement test scores*. Paper presented at the meeting of the National Council of Measurement in Education, New Orleans.

Reviewer: Ray J. Fenton.

Metropolitan Achievement Tests—7th Edition

Authors/Developers: I. H. Balow, R. C. Farr, & T. P. Hogan. The Psychological Corp., 555 Academic Ct, San Antonio TX 78204-2498.

Source: The Psychological Corporation and Harcourt Brace Jovanovich, Inc. (1992).

Cost: Classroom sets of 25 assessments for the Preprimer, $52; Primary levels $62.50. Additional cost for scoring, reporting results.

Intended context or behavior: Listening Comprehension and Listening Vocabulary.

Intended academic level: Listening is part of the Preprimer and Primary levels. Questions from content areas including reading and mathematics are read to students as part of the K-1.5 grade tests.

Time required: Preprimer, 1:35; Primer, 1:45; Primary 1, 3:40; Primary 2:40.

Test description: Norm-referenced, multiple-choice tests contain items organized by content and difficulty. Content is based on what the publisher describes as a "national consensus curriculum" which reflects common items of instruction that are included in text and supplemental materials used in U.S. schools. Listening skills include understanding explicit information, making inferences, following oral instructions, anticipating predictable language, and distinguishing between "real" and "make-believe" in a story.

Scoring: Tests are machine-scored; however, hand-scoring materials are available.

Norm/criterion reference populations: National norming population included over 10,000 students.

Culture or gender biases: Advisory panel of minority-group educators reviewed items for bias. Statistical analysis checked for items that discriminated among test takers on factors other than overall ability (performance).

Validity

Authors provide evidence of a match between test items and instructional domains.

Reliability

Test/retest: Greater than .80

Evaluative Reactions

Practicality and usefulness: A well-made norm-referenced test that may provide insight into the overall success of an academic program relative to a national measure. The great advantage of the listening test is that it is based on items within academic content areas. Scores may identify students who may need additional instruction in listening skills. *Classroom Communication Skills Inventory: A Speaking and Listening Checklist* (CCSI) may be used in conjunction with the MAT7 to identify students who are performing below teacher expectations for each grade level.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: The listening sections of the tests are read by teachers. The advantage of teacher administration is that students are used to listening to the reader, the disadvantage is that delivery is not standardized—teachers are only asked to read slowly and clearly. No information is provided on consistency of scores across administrators.

Primary strengths: The listening test is embedded in measures of accomplishment in various content areas.

Primary weaknesses: There is no effort to consider listening beyond the dictation of problems, sentences with embedded vocabulary items, and the like. There is no concern for understanding natural conversation.

Overall adequacy: The test is a well-developed measure of academic performance. The listening test is a useful tool that emphasizes the importance of listening in skill development of language and understanding in all content areas.

Additional Reference:

Franklin, M. R., Roach, P. B., Clary, E., Jr., & Ley, T. C. (1993). Overcoming the reading comprehension barriers of expository texts. *Educational Research Quarterly, 16*(1), 5-14.

Related References (from previous editions):

Decker, B. C. (1991). *Early literacy instruction with computers and whole language: An evaluation of the Writing-To-Read Computer Program with disadvantaged minority children.* (ERIC Document Reproduction Service No. ED 335 635)

Decker, B. C. (1991, May). *Early literacy instruction with computers and whole language: An evaluation of the Writing-To-Read Computer Program with disadvantaged minority children.* Paper presented at the meeting of the International Reading Association, Las Vegas.

Walker, K. F., & Burris, B. (1991). Correlation of Sensory Integration and Praxis test scores with Metropolitan Achievement Test scores in normal children. *Occupational Therapy Journal of Research, 11*(5), 307-310.

Reviewer: Ray J. Fenton.

Michigan Educational Assessment Program (MEAP) Listening Test

Authors/developers: Michigan Department of Education. Michigan Dept of Education, Ottawa Street Office Bldg, PO Box 30008, Lansing MI 48902, 517-373-3324.

Source: Michigan Department of Education (1986).

Cost: Review copies available upon request.

Intended context or behavior: Listening.

Intended academic level: Grades 4, 7, 10.

Time required: 45 minutes or less.

Test description: This is a multiple-choice response listening test in which students respond to stories presented on audiotape. The test examines multiple objectives related to listening in an educational environment. Students must recall details, draw conclusions, identify sequence of activities, identify purpose, differentiate fact or opinion, and infer character traits of the speaker.

Scoring: Scoring involves multiple-choice answers.

Norm/criterion reference populations: Average scores are available for samples of students drawn from the Michigan student population.

Culture or gender biases: Items do not appear to have a cultural or gender bias.

Validity

Concurrent: Scores are available showing performance on the Michigan Reading Test which show substantial agreement on four point scales of reading and listening attainment.

Face/content: Items are typical of classroom listening activities but do not reflect the ability to draw conclusions from a variety of listening situations. Stories presented come from a variety of academic areas such as science and social studies.

Construct: Test covers common critical listening objectives.

Reliability

Information not available

Evaluative Reactions

Practicality and usefulness: Easily administered for classroom screening. Listening to stories may not generalize well to listening in other settings.

Clarity of instructions: Instructions are clear and easy to follow.

Standardization in administration and evaluation: Recorded stimuli and multiple-choice answers.

Range of responses or degree of competence: Most students score above the average on Levels 3 or 4 of the four-level, Michigan achievement scale.

Primary strengths: Easy administration.

Primary weaknesses: Lack of information on reliability.

Overall adequacy: Adequate screening measure with demonstrated relationship to reading achievement.

Additional References

Runkel, P. E. (1986). *Information of the recommendations of the Superintendent's Study Group on MEAP plans for the future.* Lansing, MI: Michigan State Department of Education. (ERIC Document Reproduction Service No. ED 271 487)

Saginaw Public Schools. (1993). *Correlational study into the relationship between grade point averages, Michigan Educational Assessment Program (MEAP) scores and student absences for tenth grade Arthur Hill and Saginaw High School students, 1992-1993.* Saginaw, MI: Author.

Reviewer: Neil J. O'Leary.

National Achievement Test—2nd Edition

Authors/Developers: J. W. Wick, J. M. Mason, & J. Stewart. American College Testing, PO Box 2270, Iowa City IA 52240.

Source: American College Testing (1990).

Cost: $47 for a classroom set of materials for 25 students. Scoring is an additional cost.

Intended context or behavior: Listening.

Intended academic level: Grades K, 1, and 2.

Time required: Approximate time of the Listening Vocabulary and Listening Comprehension Tests is 20 minutes per test plus 5 minutes for distribution of materials.

Test description: As the teacher reads the items, the student selects the answers from three or four choices provided in a machine scorable test booklet. Listening Vocabulary consists of 30 sentences with a target word used in context. The student marks the space under a picture of the object identified in the sentence. Listening Comprehension consists of 25 items, each of which provides some sort of direction or information. The student marks the space under a picture of a location, object, or activity described in the sentence.

Scoring: This is a machine-scored, multiple-choice test; however, materials for hand-scoring are available.

Norm/criterion reference populations: More than 150,000 students drawn from 110 public school and 30 parochial districts.

Culture or gender biases: N/A.

Validity

Concurrent: Correlations are provided with reading and other subtests of the achievement battery.

Face/content: Items are selected to reflect common language used in school and appear to be appropriate for the grade levels tested. Content is linked to behavioral objectives. Documentation makes it possible to match test content against program goals.

Reliability

Test/retest: Greater than .80.

Internal consistency: Greater than .70.

Evaluative Reactions

Practicality and usefulness: Good integration into a language arts achievement test.

Clarity of instructions: Clear and well-written manual.

Standardization in administration and evaluation: Machine-scored, norm-referenced test. Instructions are detailed and complete.

Range of responses or degree of competence: Expected scores are provided when given in conjunction with the Developing Cognitive Abilities Test. See Conoley and Kramer (1992) for a discussion of the scores.

Primary strengths: Ease of administration.

Primary weaknesses: Lack of inclusion of visual, contextual cues and short passages.

Overall adequacy: A well-made test that is a useful extension of a basic skills achievement test.

Additional reference:

Wick, J. W., Mason, J. M., & Stewart, J. National Achievement Test (1992). In J. C. Conoley & J. J. Kramer (Eds.), *The eleventh mental measurements yearbook* (11th ed., pp. 587-591). Lincoln: Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

Reviewer: Neil J. O'Leary.

Stanford Achievement Test—8th Edition

Authors/Developers: Harcourt Brace Jovanovich, Inc., & The Psychological Corporation. The Psychological Corporation, 555 Academic Ct, San Antonio TX 78204-2498.

Source: The Psychological Corporation (1991).

Cost: $60-70 per classroom set of 35 tests and administration materials. Scoring costs are additional and depend on the scores and reports desired. The test may be hand-scored.

Intended context or behavior: Listening, Vocabulary, and Listening Comprehension.

Intended academic level: Grades K-9.

Time required: 35 minutes.

Test description: As the teacher reads sentences or brief paragraphs, the students select the correct answers in an answer booklet. For listening vocabulary, students select the word or phrase that best reflects the meaning of a word presented within the context of a sentence. Listening comprehension asks students to identify a picture that corresponds to a description, solve a riddle, draw conclusions, make inferences, state relationships, or reflect the meaning of a word or phrase presented in a brief paragraph.

Scoring: The tests are machine- or hand-scored with a key.

Norm/criterion reference populations: Over 1,000 schools and 215,000 students participated in item tryouts with a minimum of 700 students per test item. Approximately 300,000 participated in the norming sample, which was representative of the national population.

Culture or gender biases: Items were reviewed by a culturally diverse panel; effort was made to balance gender references. Rasch model and Angoff delta estimates of difficulty were computed for gender and African American, Hispanic, White students.

Validity

Face/content: The publisher argues that the listening tests reflect important in-school and out-of-school skills. Examination of the prompts read by the teacher shows that they appear to be appropriate for classroom listening. There is no effort to demonstrate the ability to make inferences about the emotional state of the speaker, the motivations of the speaker, or the meaning of statements as part of an ongoing dialogue.

Construct: No analysis by cluster of like items is provided.

Reliability

Alternate forms: Limited information is presented based on a comparison with the 1982 edition of the test. Almost all correlations are .80 or higher.

Internal consistency: KR-20 of .80 or higher for listening subtests.

Evaluative Reactions

Practicality and usefulness: This appears to be a useful listening test that examines a wider range of interpretive skills than some others.

Clarity of instructions: Clear and easy to follow.

Standardization in administration and evaluation: The test is read by teachers. The advantage is that students are familiar with the person giving the test. The disadvantage is that delivery is not standardized—teachers are asked to read slowly and clearly. No information is provided on the consistency of scores across administrators.

Primary strengths: Ease of administration as part of a traditional norm-referenced test battery.

Additional References

Deeter, T., & Wilson, M. (1993). *Assessment program results 1992-1993*. Des Moines: Des Moines Public Schools. (ERIC Document Reproduction Service No. ED 370 972)

Fremer, J. (1991). *Edumetric considerations in the design of the new SAT*. (ERIC Document Reproduction Service No. ED 339 717)

Fremer, J. (1991, August). *Edumetric considerations in the design of the new SAT*. Paper presented at the meeting of the American Psychological Association, San Francisco.

Harcourt Brace Jovanovich, Inc. (1992). Stanford Achievement Test (8th ed.). In J. C. Conoley & J. J. Kramer (Eds.), *The eleventh mental measurements yearbook* (11th ed., pp. 859-865). Lincoln: Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

Nelson, D. (1994). *Accountability reports for all districts and schools. The Utah statewide testing program—1993. Appendix to the Annual Report of the State Superintendent of Public Instruction*. Salt Lake City: Utah State Office of Education. (ERIC Document Reproduction Service No. ED 373 077)

Walthall, J. E., & Walthall, C. (1993). *A comparison of Stanford Achievement Test results among regular class and special education students*. Phoenix: Arizona State Department of Education. (ERIC Document Reproduction Service No. ED 370 322)

Walthall, J. E., & Walthall, C. (1993, November). *A comparison of Stanford Achievement Test results among regular class and special education students*. Paper presented at the meeting of the Mid-South Educational Research Association, New Orleans.

Reviewer: Neil J. O'Leary.

Survey of Early Childhood Abilities
(Test of Auditory Perceptual Skills)

Author/Developer: K. G. Codding. Psychological and Educational Publications, Inc., c/o Karen Gardner Codding, 1477 Rollins Rd, Burlingame CA 94010.

Source: Psychological and Educational Publications, Inc. (1987).

Cost: Manual, 25 Profile Sheets, 25 Record Booklets, 25 Test booklets, and Test Plates is $79.95 Specimen set for review is $16.

Intended context or behavior: Listening.

Intended academic level: Ages 4-7.

Time required: Tests are not timed; administration time varies with age.

Test description: The Test of Auditory Perceptual Skills is a subtest of a four-test battery that may be administered alone. This test includes Auditory Number Memory, Auditory Sentence Memory, Auditory Word Memory, Auditory Interpretation of Directions, Auditory Word Discrimination, and Auditory Processing (thinking and reasoning). The examiner reads questions to the student and records number of correct responses on a test profile. Raw scores are then converted to Language Age, Scaled, Percentile Rank, and Stanine Scores.

Scoring: Student responses are judged "correct" or "incorrect." Correct responses are totaled. For the Auditory Processing test, the student is asked to respond to questions such as "Why does a rock sink?" Then the appropriate responses are judged correct.

Norm/criterion reference populations: Norms were developed based on the performance of 962 children residing within the San Francisco Bay area. Students ranged in age from 4 years through 12 years, 11 months. An individual weighting procedure was used to approximate the level of ability for each 1-year age group based on scaled scores from the Wechsler Preschool and Primary Scale of Intelligence or the Wechsler Intelligence Scale for Children, revised.

Culture or gender biases: No significant difference was found between male and female performance at various age levels.

Validity

Predictive: Overall SECA scores correlated significantly with other achievement test scores.

Face/content: Items reflect common auditory-perceptual skills needed as part of instruction.

Reliability

Information not available

Evaluative Reactions

Practicality and usefulness: Useful instrument for screening of students for auditory understanding.

Clarity of instructions: Instructions are clear and easy to follow.

Standardization in administration and evaluation: No information is provided on standardization of "acceptable" answers and there may be some effects from nonstandard English on some of the items.

Primary strengths: Easily administered early childhood achievement test with unusually clear scoring and administration directions.

Primary weaknesses: Lack of information about the performance of speakers who use nonstandard language, interrater reliability on items where judgment is necessary, and limited tasks.

Overall adequacy: The four tests of the Survey of Early Childhood Abilities provide a good overview of student academic skills. The listening component of the test allows a consideration of the ability of the child relative to a number of important auditory skills. The test would be useful for both screening/diagnosis and as a research instrument.

Additional References

Codding, K. G. (1993). *SECA: Survey of early childhood abilities manual.* San Francisco: Children's Hospital of San Francisco.

Gardner, M. F. (1985). *Test of auditory-perceptual skills.* San Francisco: Children's Hospital of San Francisco.

Reviewer: Ray J. Fenton.

Watson Barker Listening Test—High School Version

Authors/Developers: K. W. Watson, L. L. Barker, & C. V. Roberts. Spectra, Inc., PO Box 13591, New Orleans LA 70185-3591.

Source: Spectra, Inc. (1989).

Cost: Includes videotapes and 25 test forms, $119.95.

Intended context or behavior: Listening.

Intended academic level: High School.

Time required: Approximately 35 minutes.

Test description: This is a five-part listening test presented by either audio- or videotape, including evaluation of message content, understanding meaning in conversations, understanding and remembering lectures, evaluating emotional meanings in messages, and following directions. Conversational segments focus on content and interactions familiar to high-school students.

Scoring: The multiple-choice responses are hand-scored by teacher or test takers.

Norm/criterion reference populations: Average scores for 137 males and 183 females are presented in the Facilitator's Guide. General scoring guidelines are presented for characterizing scores from "Very Poor" to "Excellent."

Culture or gender biases: Male and female, black, white, and Asian speakers are included in the source materials. Females score slightly higher on the measure than males.

Validity

Face/content: Situations appear to be common to high-school students.

Construct: The instrument has been used by a number of communication researchers.

Reliability

Alternate forms: Form A and Form B show a significant but low (.53) correlation. Correlations of subsections range from .11 to .38.

Evaluative Reactions

Practicality and usefulness: A very practical and easy to administer test that is consistent with various methods of teaching listening skills. Instrument focuses on listening in typical communication situations.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Tape and video versions guarantee standard presentation of material.

Primary strengths: The test is easy to administer and is a valuable instructional tool. The test has been widely used for research and instruction.

Primary weaknesses: Information on validity and reliability presented in the manual is limited. Additional information is available from the authors.

Overall adequacy: Well-balanced situations and ease of administration make this an excellent choice for secondary classrooms.

Additional References

Karr, M., & Vogelsang, R. W. (1989). *A comparison of the audio and pilot video versions of the Watson-Barker Listening Test: Forms A & B*. Portland, OR: Portland State University. (ERIC Document Reproduction Service No. ED 304 729)

Karr, M., & Vogelsang, R. W. (1989, February). *A comparison of the audio and pilot video versions of the Watson-Barker Listening Test: Forms A & B*. Paper presented at the meeting of the Western States Speech Communication Association, Spokane.

Roberts, C. V. (1988). The validation of listening tests: Cutting the Gordian Knot. *Journal of the International Listening Association, 2*, 1-19.

Rubin, R. B., & Roberts, C. V. (1987). Comparative examination and analysis of three listening tests. *Communication Education, 36*, 142-153.

Rubin, R. B., & Shepherd, P. (1985). *Listening assessment via the Communication Competency Assessment Instrument*. Kent, OH: Kent State University. (ERIC Document Reproduction Service No. ED 264 616)

Rubin, R. B., & Shepherd, P. (1985, November). *Listening assessment via the Communication Competency Assessment Instrument*. Paper presented at the annual meeting of the Speech Communication Association, Denver, CO.

Watson, K. W., & Barker, L. L. (1988). Listening assessment: Validation of the Watson-Barker Listening Test. *Journal of the International Listening Association Convention, 2*, 20-32.

Watson, K. W., & Barker, L. L. (1992). Watson Barker Listening Test—High School Version. In J. C. Conoley & J. J. Kramer (Eds.), *The eleventh mental measurements yearbook* (11th ed., pp. 1,009-1,012). Lincoln: The Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

Watson, K. W., Barker, L. L., & Roberts, C. V. (1989). Watson Barker Listening Test. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (10th ed., pp. 384-385). Lincoln: The Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

Reviewer: Ray J. Fenton.

Diagnostic Analysis of Nonverbal Accuracy Scale

Author/Developer: S. Nowicki, Dept of Psychology, Emory Univ, Atlanta GA 30301.

Source: Stephen Nowicki (1992).

Cost: Contact authors. Slides and other test materials are available for duplication cost.

Intended context or behavior: Nonverbal.

Intended academic level: Grades 1-5, Ages 6-10.

Time required: Receptive tests may be group administered in less than 30 minutes, Expressive tests must be individually administered in less than 10 minutes.

Text description: Scoring: Objective scoring of subject's identification of the proper emotion is performed on the receptive tests. Observer/listener identification of the accuracy of the emotion represented is performed on the expressive tests. The expressive tests are video- or audiotaped. The scorers work from tapes to allow determination of interrater reliability.

Norm/criterion reference populations: Age and grade level norms are available from the author. Norms are derived from experimental populations and, while care has been taken in the description of populations, some caution should be used in making interpretations.

Culture or gender biases: No significant differences by gender have been found.

Validity

Predictive: Studies show correlations with popularity ratings, Piers-Haris self-image, and academic achievement, but not with IQ.

Face/content: Joint administration with the PONS for a small number of students showed a correlation of .53.

Reliability

Scoring: Trained raters have demonstrated greater than .90 agreement.

Evaluative Reactions

Practicality and usefulness: Instrument is useful for screening of students. More research is needed to establish stable grade and age level norms.

Clarity of instructions: Acceptable.

Standardization in administration and evaluation: Directions are clear for administration. More information would be helpful in the preparation of raters for evaluating student responses.

Range of responses or degree of competence: Both receptive and expressive tests appear to provide adequate measure of the ability to identify or portray the four emotional states—happy, sad, angry, and fearful.

Primary strengths: This innovative test combines standard nonverbal assessment methods for children.

Primary weaknesses: Limited information is available on various cultural-ethnic group differences. The instrument does not examine the relationship between the receptive and expressive tests and drawing conclusions from actual interpersonal communication situations.

Overall adequacy: Valuable tool for research with potential for classroom applications. The test should prove useful as a screening measure or as part of an instructional program. It has been used in a number of research studies.

Additional References

Nowicki, S., Jr. (1990). *The relation of the DANVA to the PONS.* Unpublished manuscript, Emory University, Atlanta.

Nowicki, S., Jr., & DiGirolamo, A. (1989). The association of external locus of control, nonverbal processing difficulties and emotional disturbance. *Behavioral Disorders, 15,* 28-34.

Nowicki, S., Jr., & Duke, M. P. (1992). The association of children's nonverbal decoding abilities with their popularity, locus of control, and academic achievement. *Journal of Genetic Psychology, 153,* 385-394.

Nowicki, S., Jr., & Duke, M. P. (1993). *Individual differences in the nonverbal communication of affect: The diagnostic analysis of Nonverbal Accuracy Scale.* Unpublished manuscript, Emory University, Atlanta.

Reviewer: Ray J. Fenton

Test of Pragmatic Skills—Revised Edition

Author/Developer: B. B. Shulman. Communication Skill Builders, Inc., 3830 E Bellevue, PO Box 42050, Tucson AZ 85733. 602-323-7500.

Source: Communication Skill Builders (1986).

Cost: $65.

Intended context or behavior: Listening and Interpersonal Communication.

Intended academic level: Ages 3- 8 years, 11 months.

Test description: This test is individually administered and based around four activities: playing with puppets, pencil and paper, telephones, and blocks. There are 34 standard probes which may be repeated and relate to a variety of skills including: greeting, answering, informing, naming, requesting information, rejection-denial, requesting action, and closing conversation. An informal and non-normed Language Sampling Supplement is available which includes conversational turn taking, speaker dominance, topic maintenance, and topic change.

Scoring: Model responses are suggested. The examiner rates each response on a 0-5 scale: (0) no response, (1) contextually inappropriate response, (2) contextually appropriate nonverbal/gestural response only, (3) contextually appropriate one-word response without elaboration, (4) contextually appropriate response with minimal elaboration (2-3 words), and (5) contextually appropriate response with extensive elaboration (more than 3 words). Under ideal conditions, the examiner may videotape the responses and complete scoring after the test administration session. The subject's score is not based solely on the number of words uttered, but also on test administrator's rating of appropriateness of the response for the context.

Norm/criterion reference populations: Normative data sheets illustrate mean scores for performance of various age groups. Standardization took place with 650 Anglo middle class students between ages 3 years and 8 years 11 months who had been tested for normal auditory comprehension. There were an equal number of male and female students who were drawn from the Northeast, North Central, South, and West. Percentile ranks are provided.

Culture or gender biases: None are included. It appears that nonstandard speech patterns might bias ratings. No information is reported on trials with non-Anglo students.

Validity

Concurrent: Clinicians provided "overall" assessments of two standardization subjects which correlated .68 and .64 with the Level of Response Rating Scale.

Face/content: Exercises and probes appear to be representative of typical conversation in a variety of contexts.

Construct: Underlying constructs represented by the rating scales relate to important communication skills commonly assessed in primary school.

Reliability

Test/retest: Three-week delay test/retest of 120 students by the same examiner resulted in a Pearson Product Moment correlation of .96.

Scoring: Evaluative Reactions

Practicality and usefulness: Easily administered individual assessment that provides information on the communication skills of young students. All interactions are adult/child and stimulated by adult initiated probes.

Clarity of instructions: Instructions are clear and easy to follow.

Standardization in administration and evaluation: Directions for administration and scoring are clear. No clear benchmarks are provided for expected performances of students at various age levels or from various cultural backgrounds. Additional information is needed on the relationships between various exercises and between various raters.

Range of responses or degree of competence: Ratings are based on a normal population and scores do not support classification of students as deficient or in need of special, remedial instruction. It appears that the average score expected on the individual tasks is quite high by the time students reach 8 years 11 months.

Primary strengths: Clear directions, innovative situations, and validation make this one of the best instruments for assessment of communication skills in the early grades.

Primary weaknesses: Lack of validation with non-Anglo students and lack of information on the relationship among the activities.

Overall adequacy: This is a good instrument for use in either the classroom or as a research instrument for normal students.

Additional References

Shulman, B. B. (1984a). *Using play behavior to describe young children's conversational abilities*. Huntsville: University of Alabama. (ERIC Document Reproduction Service No. ED 264 019)

Shulman, B. B. (1984b). *Using play behavior to describe young children's conversational abilities*. Paper presented at the meeting of the International Congress for the Study of Child Language, Austin.

Shulman, B. B. (1984c). *Using play behavior to describe young children's conversational abilities*. Paper presented at the meeting of the National Association for the Education of Young Children, Los Angeles.

Shulman, B. B. (1989). Test of Pragmatic Skills. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (10th ed., pp. 847-850). Lincoln: The Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

Reviewer: Neil J. O'Leary.

BEST COPY AVAILABLE

Oral Proficiency Examination

Authors/Developers: Gary Community Schools Corporation. George Comer, Director of Curriculum Services, Gary Community Schools Corporation, 1234 Cleveland Street, Gary IN 46401; 219-977-2171.

Source: Gary Community Schools Corporation (1977).

Cost: Free/public domain.

Intended context or behavior: High School Oral Proficiency Exam includes interview or short presentation.

Intended academic level: Grade 10—Administered as part of a Grade 10 general speech class.

Time required: Variable but fewer than 10 minutes per students.

Test description: Students participate in an interview with a member of a three-person assessment committee. The interview is taped. Three raters score student responses to three or four open-ended questions selected by the interviewer. Raters rate each student on a 4-point scale: (1) severely deficient, (2) deficient to moderately deficient, (3) average proficiency, and (4) moderate to high proficiency. Students are rated on articulation, pronunciation, verbal utterances, rate, word usage, voice qualities, volume, and sentence structure. Raters use a "holistic" scale based on examples of performance. Raters review these typical examples of behavior in advance to improve the consistency of ratings, and this review process acts as the standards for rating. Students may be rescored if there are discrepancies among ratings. A tape recording of the interview becomes a part of the student's permanent record.

Scoring: Trained speech teachers rate students on "holistic" scales after training with benchmark examples.

Norm/criterion reference populations: Benchmark examples are drawn from 10th grade speakers.

Culture or gender biases: None are indicated. Without examples of the benchmarks used for scoring, there is a question about how the test would be applied to students who are not speakers of standard English.

Validity

Face/content: Activity mirrors other classroom interview activities.

Construct: Eight scales used in the test directly relate to the K-12 speech goals of the Gary School Corporation and the content of required courses.

Reliability

Scoring: Documents indicate that raters are frequently checked for wide discrepancies in scoring. Raters only achieve adequate interrater reliability if appropriate training is provided.

Internal consistency: This appears to be a potential problem since not all students are asked to respond to the same stimuli.

Evaluative Reactions

Practicality and usefulness: This proficiency exam is easily administered and is much like many informal measures suggested with the goal-based curriculum approach. It is included here because of the detail provided in the discussion of scoring and the use of benchmarks in scoring. There is limited usefulness because students are only asked to answer three or four questions as an indicator of overall proficiency.

Clarity of instructions: Instructions and rating scale are clear.

Standardization in administration and evaluation: Benchmark samples were not provided. A formal procedure for checking interrater reliability should be incorporated into the testing procedure.

Primary strengths: Easily administered screening test to judge elements of voice and speaking.

Primary weaknesses: Lack of information on validity and reliability and the limited scope of the assessment.

Overall adequacy: This may prove useful as a general screening exam but is limited as an indicator of proficiency in communication.

Additional Reference:

Gary Community School Corporation. (1978). *Oral Proficiency Program*. Gary, IN: Author.

Reviewer: Ray J. Fenton.

Communication Competency Assessment Instrument—High School

Author/Developer: R. B. Rubin. Spectra, Inc. Publishers, c/o Barbara Nixon, 4197 Monterey Dr., Memphis TN 38128. 901-386-7666.

Source: Spectra, Inc. (1993).

Cost: $39.95 for Rating Manual.

Intended context or behavior: Speaking/interaction. A similar instrument for college students includes a listening component.

Intended academic level: Grades 9-12.

Time required: Manual suggests less than 15 minutes per individual tested. Actual test time for the student is as much as 30 minutes including time to prepare a speech.

Test description: Oral communication assessment is based on competencies identified by SCA as important for success in school. Students present a 3-minute impromptu talk and engage in a guided conversation with teacher/rater. Skills assessed include use of understandable pronunciation, appropriate tone of voice, clarity, distinguishing between informative and persuasive messages, presenting thesis and main points clearly and concisely, expressing and defending a point of view, recognizing when others do not understand, introducing yourself appropriately, obtaining information about a career, answering instructor's question about classroom performance, expressing satisfaction or dissatisfaction to an instructor about a course, using chronological order to explain your activities throughout the day, giving directions, describing the viewpoint of a fellow student who disagrees with your evaluation of a class, and describing differences in opinion about steps necessary to accomplish academic or vocational goals.

Scoring: Students are scored on 5-point criterion referenced scales: 3 denotes acceptable behavior, 5 represents "superior competency," and 1 shows lack of competency.

Norm/criterion reference populations: The education context of the competencies is based on the Speech Communication Association's *Speaking and Listening Competencies for High School Graduates*.

Culture or gender biases: A review of the highly similar CCAI by a panel of racially-diverse speech communication professionals found no group-specific language or vocabulary. CCAI performance by a group of college students found no significant difference in attaining a satisfactory score (3.00) by gender or minority group membership.

Validity

Face/content: Individual competencies are related to important skills for success, identified by SCA. Speech communication faculty members were able to sort individual questions into the main SCA competencies.

Reliability

Scoring: Four graduate students attained .80 interrater reliability after two hours of training. On the longer college version, raters have attained reliabilities over .90.

Internal consistency: Alpha reliability for the CCAI-HS was .77; split-half was .66.

Evaluative Reactions

Practicality and usefulness: Highly useful direct measure of student competence on critical communication skills that are important for school success.

Clarity of instructions: Instructions are clear and easy to follow.

Standardization in administration and evaluation: Training materials are adequate to prepare raters. Administration instructions are clear.

Range of responses or degree of competence: Scoring categories have high face validity.

Additional References

Rubin, R. B. (1994). The CCAI-HS: A new assessment tool. *The Speech Communication Teacher, 8*(2), 14-15.

Rubin, R. B. (1985). Validity of the Communication Competency Assessment Instrument. *Communication Monographs, 52*, 73-185.

Rubin, R. B. (1982). Assessing speaking and listening competence at the college level: The Communication Competency Assessment Instrument. *Communication Education, 31*, 19-32.

Rubin, R. B., Welch, S. A., & Buerkel, R. (1994, November). *Assessing high school speech instructional outcomes*. Paper presented at the meeting of the Speech Communication Association, New Orleans.

Rubin, R. B., Welch, S. A., & Buerkel, R. (1995). Performance based assessment of high school speech instruction. *Communication Education, 44*, 30-39.

Reviewer: Neil J. O'Leary.

Hunter-Grundin Literacy Profiles

Authors/Developers: E. Hunter-Grundin, & H. U. Grundin. The Test Agency, Cournswood House, North Dean High St, Wycombe, Bucks, England. (44-24-024) 3384.

Source: The Test Agency.

Cost: Available from publisher. Price is quoted in British pounds.

Intended context or behavior: Spoken language is included with a Profile of Personal Interests and other measures of Reading for Meaning, Spelling, and Free Writing.

Intended academic level: Levels 1-5 which correspond to various age ranges: I = 6.5-8.0; II = 7.1-9.3; III = 9.0-10.0; IV = 9.1-11.5; V = 10.1-12.7.

Time required: Times vary by level but are generally 3-10 minutes per assessment. Speaking assessment is individually administered and takes about 3 minutes for each student.

Test description: Students are asked to look at and "tell me what is going on" in a complex picture showing a typical school lunchroom scene. Student discourse is scored on five Spoken Language Scales: Confidence, Enunciation, Vocabulary, Accuracy, and Imagination.

Scoring: The teacher grades each attribute "A through E", highest to lowest, based on rubric-like descriptions of specific behaviors which are representative of each level. For example, Confidence is scaled from A, "Highly confident, totally relaxed" to E "Obvious lack of confidence, high degree of anxiety/nervousness."

Norm/criterion referencing: Measure has been used with 2,500 or more English and Scottish students at each test level.

Culture or gender biases: N/A.

Validity

Face/content: Activity is typical of classroom activity and would be natural for the student being assessed.

Reliability

Information not available

Evaluative Reactions

Practicality and usefulness: The authors claim that the assessment provides teachers with "a simple and practical means of keeping records of the progress of individual children" in language development. The manual includes teachers' testimonial statements that indicate a high degree of student satisfaction and enjoyment.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Administration is straightforward, but no examples are given of actual performances that could be used to ground the teacher's interpretation of student utterances relative to the various scales.

Range of responses or degree of competence: Recommended scaling and criteria for classification appear to be adequate.

Primary strengths: The Hunter-Grundin might serve as a model for an integrated language arts assessment that includes oral communication. The Spoken Language Scales are easy to administer and take only 3 to 5 minutes per student. The scoring scales are well thought out and provide clear distinctions between levels of achievement.

Primary weaknesses: No information on reliability or validity is provided for spoken language use. Reliability and validity information is quite high for other sections of the tool. No information is available that would allow normative comparisons with students of similar age or level.

Overall adequacy: The measure is adequate for classroom assessment and would be useful for research and group comparisons if validity and reliability were established.

Additional References

Quigley, M. A. V. (1993a). *The improvement of reading comprehension skills in at risk second graders.* Fort Lauderdale, FL: Nova University. (ERIC Document Reproduction Service No. ED 359 505)

Quigley, M. A. V. (1993b). *The improvement of reading comprehension skills in at risk second graders.* Unpublished master's thesis, Nova University, Fort Lauderdale, FL.

Reviewer: Ray J. Fenton

Joliet 3-Minute Speech and Language Screen—
Revised Edition

Authors/Developers: M. C. Kinzler, & C. C. Johnson. Communication Skill Builders, 3830 E Bellevue, PO Box 42050, Tucson AZ 85733. 602-323-7500.

Source: Communication Skill Builders (1993).

Cost: $55.

Intended context or behavior: Phonomological, grammatical and semantic screening of normal students.

Intended academic level: Grades K, 2, and 5.

Time required: 3 minutes per student.

Text description: This is an individually administered screening measure of vocabulary, grammar, phonology, voice, and fluency designed to identify students who may need additional assessment and review for speech and language services. The test includes eight vocabulary plates for each grade (24) and 10 evoked sentences. Students are asked to point to illustrations of objects for the listening/vocabulary assessment and to repeat 10 sentences. The teacher or Speech and Hearing Specialist assesses each student relative to specific grade level criteria and identifies students falling below the criterion score as deficient.

Scoring: A scoring sheet is provided, and errors are recorded for each tested item. Some sentences allow for identification of up to three specific phonological errors. Error scores are compared to suggested cutoff scores. Directions indicate that students may be retested.

Norm/criterion reference populations: The test was standardized on 2,587 students in grades K, 2, and 5 in eight public and parochial schools in Joliet, Illinois. A randomly selected subsample of 586 was used to establish cutoff scores. Students were screened to assure that there were no hearing disorders and that they were mainstreamed for the greater part of the school day.

Culture or gender biases: Some differences were found. There is extended discussion of nonstandard dialects, but only some discussion of acceptable and unacceptable deviations from common pronunciation.

Validity

Predictive: Pass/Fail categorizations were compared to mean differences in PPVT and Carrow Elicited Language Inventory scores for grade and ethnic groups. Gender was found to be a significant variable related to vocabulary but not grammar. Within racial/ethnic classifications, substantial differences were found between Hispanic and White groups. Within economic groups, low income groups were significantly different from high/mid groups.

Face/content: Activities are clearly related to school and home communication behaviors. There is emphasis on interpretation and generation of speech.

Reliability

Test/retest: No significant difference was found in the performance of students over a 4-week period.

Scoring: Criterion-based scoring appears to be straightforward. No interrater reliability is provided.

Evaluative Reactions

Practicality and usefulness: Highly useful for screening students in early grades. The manual provides the basis for a discussion of fluency and listening vocabulary with parents.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Additional information on interrater reliability would be helpful.

Range of responses or degree of competence: Recommended criteria for classification appear to be adequate.

Primary strengths: Ease of administration and high face validity.

Primary weaknesses: Lack of information on interrater reliability and correlation with other communication skills.

Overall adequacy: A useful measure for the early-grades classroom teacher interested in a quick measure of oral language skills in the classroom. A very useful instrument for initial screening of students for speech and language services or determining overall communication proficiency of young students.

Additional References: Not available

Reviewer: Ray J. Fenton

Oral Communication Assessment Instruments For Higher Education

Ellen A. Hay, *Augustana College*

For colleges and universities, the pressure to demonstrate the effectiveness of instruction has increased significantly in recent years. A 1990 survey indicated that 82 percent of American colleges and universities are engaged in some form of assessment (El-Khawas, 1990). The motivation for this increased pressure stems from a variety of forces. First, increased accountability is demanded by legislatures and accrediting agencies. Institutions are now required to develop clearly articulated goals and objectives and then to offer proof that such goals and objectives are accomplished. Assessment has become the means for determining whether we are reaching our stated outcomes. Second, as a result of extensive criticism of higher education and increasing competition for students, more emphasis is placed on improving instruction at the college level. Classroom assessment helps educators make decisions about what will and will not be included in particular courses. Overall, assessment is viewed as a critical component in this improve-ment process, for it informs faculty and administrators of what is and what is not working.

Due to these pressures and opportunities, planning and implementing assessment activities have become a priority at many institutions. The primary emphasis has been in two areas. Faculty and administrators have concentrated mainly upon the evaluation and improvement of the general education curriculum and major programs of study. For departments of speech communication, this has meant focusing upon questions such as:

◊ Why should students be required to complete coursework in speech communication?

◊ How can communication assessment tools support effective classroom instruction?

◊ How can assessment tools inform students about their own communication behavior?

◊ How can assessment be used as feedback to the individual student?

◊ What are the expected outcomes of required course work in the field?

◊ How can the outcomes for a required basic or advanced speech course best be demonstrated?

◊ What are the goals and objectives of the speech communication major?

◊ How do these goals and objectives relate to the institutional mission?

◊ How successful is the instruction designed to reach these outcomes?

◊ What means are available to demonstrate the outcomes of instruction?

◊ How can instruction be designed to meet the needs of students?

Departments attempting to answer these questions face the challenge of either designing or assembling assessment instruments for the dual (and not mutually exclusive) purposes of: (a) demonstrating the effectiveness of instruction by measuring student achievement, and (b) providing effective information and feedback to the students regarding their communication behavior. The review of instruments provided herein is designed to assist departments and instructors with both purposes. It includes descriptions of 45 instruments that are divided into two categories based on these purposes. The first section contains instruments intended for skills and behavioral assessment. The second section contains instructional and informative measures. Undoubtedly no one instrument will be suitable for all contingencies; but, using multiple measures, linked to departmental objectives and instructional practices over an extended period, can provide valuable information on the outcomes of instruction in speech communication. What follows is a brief description of the process by which instruments were identified and then selected for inclusion in this compendium. Then a few ideas for using the instruments are provided.

Identifying and Selecting Measures

The first task in preparing this section was to identify possible instruments for review. A number of sources were consulted. Several compendiums of measurement instruments were especially valuable. These included *Measures of Speech Communication* (Educational Testing Service, 1991), *Hand-*

book of Interpersonal Competence Research (Spitzberg & Cupach, 1988), *Handbook for the Study of Human Communication* (Tardy, 1988), and *Studying Interpersonal Interaction* (Montgomery & Duck, 1991). To ascertain what measures were currently used in both research and instructional settings, issues from the last 5 years of *Communication Education* and *Communication Monographs* were reviewed, and an ERIC search was completed. A letter was sent to each state official responsible for higher education, asking about assessment practices in that state. An announcement was also placed in *Spectra* detailing the project and requesting measures.

Hundreds of measures were identified from the aforementioned search. The next task was to select those instruments most appropriate to this publication. In making these decisions, the following factors played a part:

◊ Attention to the SCA criteria guidelines for assessment.

◊ Suitability/adaptability to the instructional setting as well as for research.

◊ Appropriateness to a college/university population as opposed to more specialized populations in health care, family counseling, education, business, etc.

◊ Availability of measure from a publisher, document service, or through the library of a major research institution.

◊ Recency of measure reflecting contemporary research on the construct/concept being assessed.

Every effort was made to insure the inclusion of measures that would tap a broad spectrum of the communication discipline and also would be suited to many different types of students in a variety of educational settings. Some measures suitable for organizational communication purposes have been selected for review; those chosen, however, seem to have a wider application than just the organizational context. For educators seeking instruments that are more specifically tailored to the organizational setting, a review by Shockley-Zalabak and Hulbert-Johnson (1993) provides comprehensive data on over 60 instruments. A review by Beebe and Barge (1993) discusses evaluation forms applicable to the group setting. Since many of these instruments do not contain adequate information on validity and reliability, they have not been included in this publication, but they could be adapted and tested for use in the organizational and group communication contexts.

Decisions regarding the inclusion of measures also considered the three domains of oral competence (knowledge, skill, motivation), with knowledge only minimally and inferentially represented in the compendium. By assessing skills, one can make inferences about knowledge (and perhaps motivation). Skills and motivation are well-represented in the compendium, albeit frequently with self-report measures. Despite the fact that self-perceptions are only valid measures of internal states, such as attitudes and thoughts, the self-report tools for skills and behaviors are included to be used instructionally.

Finally, decisions were made advisedly regarding the inclusion of measures of trait and style. Such tools were included despite the notions that (a) traits are difficult to alter, and (b) styles are neither right or wrong, good or bad.

Instructors may find the use of these measures valuable in the classroom setting to inform students regarding their predispositions to act in certain ways. The constructs assessed when using these measures are related to communication competency, but they don't specifically constitute one of its three domains. The tables that precede the compendium of measures clearly indicate the type, purpose, and mode of each instrument. Using that information, the instructor or administrator can make an informed choice regarding the value and suitability of the tool for his/her classroom, department, or institution.

Using Measures

With the increased vigor of the assessment movement, criticisms of the process have also emerged. Concerns about the relationship between the curriculum and the test, and fears about the role of assessment in the faculty evaluation process are just a few of the reservations that are expressed. In selecting possible measures and developing assessment programs, it is important to be sensitive to such concerns. A working group at the SCA-sponsored 1990 Denver Conference on Oral Communication Assessment identified several assessment issues and devised a series of guidelines to inform the process (Hay, 1992a; Hay, 1992b). These guidelines are designed to insure that assessment is a positive experience intended to improve and enhance speech communication instruction.

To insure that the test matches the curriculum, a number of departments have elected to design their own tests and/or performance measures rather than rely on outside sources. Such programs start with a definition of standards appropriate to the department. Department faculty determine what students should know and what they should be able to do at the end of a particular course or program of study. The faculty then develop approaches such as paper-pencil tests, performance measures, portfolio assignments, capstone courses, and/or senior projects or theses to ascertain how well students are meeting the departmental standards. *Program Assessment in Speech Communication* (Hay, 1992b) includes several examples of colleges and universities that have gone through this process. The University of Northern Colorado (Arneson, 1993), the University of Wisconsin-Oshkosh (Willmington, 1992), and the University of Missouri-Kansas City (Aitken & Neer, 1992) also provide models for departments wishing to pursue this approach.

The compendium of instruments that follows includes some measures that are suitable for assessment when it is intended to verify the outcomes of instruction, and some instruments that can be used in instruction to provide more information about students, to start discussion, or to illustrate particular traits or dispositions. The compendium is organized around contexts or categories in communication. Instruments have been grouped by their primary focus into one of the following sections: listening competence, interpersonal competence, public speaking competence, communication apprehension, conflict management, or various other dimensions of competence. This arbitrary organization certainly should not limit use of a particular instrument for another context. When a particular measure is deemed suitable, the author or publisher of the measure should be contacted for additional information and permission for its use.

Whether assessment occurs with departmentally developed instruments or with instruments presented in this compendium, it is important to attend to the practices of effective assessment. Part I of this volume details those practices. Used appropriately, assessment can be seen as a positive force that informs the instructional process and validates the outcomes of study in speech communication.

References

Arneson, P. (1993). *Authentic exhibition assessment for speech communication*. Unpublished manuscript. Boulder, CO: University of Northern Colorado.

Aitken, J. E., & Neer, M. (1992). A faculty program of assessment for a college level competency-based communicator core curriculum. *Communication Education, 41*, 270-286.

Beebe, S. A., & Barge, J. K. (1993, November). *The competent small group communicator: A search for competencies and assessment tools*. Paper presented at the annual meeting of the Speech Communication Association, Miami.

Educational Testing Service (1991). *Measures of speech/communication*. Princeton, N.J.: Author.

El-Khawas, E. (1990). *Campus trends, 1990*. Washington, DC: American Council on Education.

Hay, E. A. (1992a). A national survey of assessment trends in communication departments. *Communication Education, 41*, 247-257.

Hay, E. A. (1992b). *Program assessment in speech communication*. Annandale, VA: Association for Communication Administration.

Montgomery, B. M., & Duck, S. (1991). *Studying interpersonal interaction*. New York: Guilford Press.

Shockley-Zalabak, P., & Hulbert-Johnson, R. (1993, November). *Oral communication competency and its assessment in organizations:* presented at the meeting of Speech Communication Association, Miami.

Spitzberg B. H., & Cupach, W. R. (1988). *Handbook of interpersonal competence research*. New York: Springer-Verlag.

Tardy, C. H. (1988). *Handbook for the study of human communication*. Norwood, NJ: Ablex.

Willmington, S. C. (1992, October). *The validity and reliability of a performance test of interpersonal communication proficiency*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Editors' Note

In the following reviews of instruments for higher education, the authors strived for consistency and uniformity in the presentation of each measure/assessment tool. However, variability exists across the instruments in terms of what information is and is not accessible and available. For example, not every instrument has been subjected to exactly the same testing procedures so not the same testing information is available for each measure. In some cases, cost or other information is not available without the user contacting the developer regarding his or her precise user needs.

Considering this variability, the reviewers chose to include information that was available and to indicate "not available" after any item to which they had no access. If users of this volume have information that could enrich the presentation of any assessment tool, it should be sent to the author, Ellen A. Hay.

Skill and Behavioral Oral Communication Assessment Instruments: Higher Education

Title	Source	Tested Context, Behaviors	Target Populations	Type/Purpose*	Mode Of Administration	Page
Listening Instruments						
Kentucky Comprehensive Listening Test	Bostrom & Waldhart	listening	college, adult	skills/behaviors	subjects listen to audiotape and respond on an answer sheet	51
Steinbrecher-Willmington Listening Test	U of Wisconsin-Oshkosh Dept of Communication	listening	college	skills/behaviors	subjects view videotape, respond on multiple-choice questionnaire	52
Watson-Barker Listening Test	Watson, Barker, & Roberts (1989)	listening	adults	skills/behaviors	subjects are presented audio- or videotape, respond in test book	53
Interpersonal Instruments						
Conversational Skills Rating Scale	Spitzberg (1995)	interpersonal	college, adult	skills/behaviors	paper & pencil self-report or trained rater evaluates 2-person interaction	54
Public Speaking Instruments						
The Competent Speaker	Speech Communication Association (1993)	public speaking	college	skills/behaviors	subjects present a speech, tester evaluates on standardized form	55
Dantes Principles of Public Speaking	Educational Testing Service	public speaking	college, adult	cognitive, skills/behaviors	multiple-choice cognition test; subjects present a speech which is audio-taped	57
Speaking Skills Assessment College Outcome Measures	American College Testing	public speaking	college	skills/behaviors	subjects present 3 short speeches which are audio-taped	58
Communication Apprehension						
Behavioral Assessment of Speech Anxiety (BASA)	Mulac & Sherman (1974)	communication apprehension	college, adult	skills/behaviors	subjects present speeches and tester assesses anxiety on a rating form	59
Various Dimensions of Communication						
Communication Competency Assessment Instrument—Revised Edition	Spectra, Inc.	public speaking, interaction abilities, listening	college	skills/behaviors	subjects present a speech, view a videotape, then interact with a tester who scores on a rating sheet	60

* See the Editors' Foreword for an explanation of the Type/Purpose categorization system.

Instructional and Informative Oral Communication Assessment Instruments: Higher Education

Title	Source	Context/ Behaviors	Populations	Type/Purpose*	Administration	Page
Listening Instruments						
Listener Preference Profile	Watson & Barker (1993)	listening	adult	trait/instructional	paper-pencil	62
Interpersonal Instruments						
Affective Communication Test	Friedman, Prince, Riggio, & DiMatteo (1980)	interpersonal	college, adult	state/instructional	paper-pencil	63
The Argumentativeness Scale	Infante & Rancer (1982)	interpersonal	college	trait/instructional	paper-pencil	64
College Self-Expression Scale	Galassi, DeLo, Galassi, & Bastien (1974)	interpersonal	college	trait/instructional	paper-pencil	65
Conversational Sensitivity Scale	Daly, Vangelisti, & Daughton (1987)	interpersonal	college	state/instructional	paper-pencil	66
Interaction Involvement Scale	Cegala (1981)	interpersonal	college	trait/instructional	paper-pencil	67
Interpersonal Communication Competence Scale	Rubin & Martin (1994)	interpersonal	college, adult	trait/instructional	paper-pencil	68
Interpersonal Communication Satisfaction Inventory	Hecht (1978)	interpersonal	college	state/instructional	paper-pencil, recall of interaction	69
Interpersonal Problem Solving Assessment Technique	Getter & Nowinski (1981)	interpersonal	college, adult	trait/instructional	subjects give free responses to 46 interpersonal problems	70
Opener Scale	Miller, Berg, & Archer (1983)	interpersonal	college, adult	state/instructional	paper-pencil	71
Relational Communication Scale	Burgoon & Hale (1987)	interpersonal	college, adult	trait/instructional	paper-pencil after a recalled or real reaction	72
Revised Self-Disclosure Scale	Wheeless (1978)	interpersonal	college, adult	trait/instructional	paper-pencil with target person in mind	73
Speech Evaluation Instrument	Zahn & Hopper (1985)	interpersonal	college, adult	state/instructional	paper-pencil after hearing or reading a transcript	74
Verbal Aggressiveness Scale	Infante & Wigley (1986)	interpersonal	college, adult	trait/instructional	paper-pencil	75
Communication Apprehension						
Class Apprehension about Participation Scale (CAPS)	Neer (1987)	communication apprehension	college, adult	state/instructional	paper-pencil	76
Communication Anxiety Inventory	Booth-Butterfield & Gould (19860	communication apprehension	college, adult	state/instructional	paper-pencil	77
Personal Report of Communication Apprehension—24 (PRCA)	McCroskey (1982)	communication apprehension	college, adult	trait/instructional	paper-pencil	78
Receiver Apprehension Test	Wheeless (1975)	communication apprehension	college, adult	trait/instructional	paper-pencil	79

* See the Editors' Foreword for an explanation of the Type/Purpose categorization system.

BEST COPY AVAILABLE

Title	Source	Context/ Behaviors	Populations	Type/Purpose	Administration	Page
Unwillingness-to-Communicate Scale	Burgoon (1976)	communication apprehension/ avoidance	college, adult	trait/instructional	paper-pencil	80
Conflict Management						
Hall Conflict Management Survey	Hall (1986)	conflict management	college, adult	state/instructional	paper-pencil	81
Putnam-Wilson Organizational Communication Conflict Instrument (OCCI)	Putnam & Wilson (1982)	conflict management	college, adult	state/instructional	paper-pencil	82
Rahim Organizational Conflict Inventory II (ROCI-II)	Rahim (1983)	conflict management	college, adult	state/instructional	paper-pencil	83
Ross-DeWine Conflict Management Message Style (CMMS)	Ross & DeWine (1988)	conflict management	college, adult	trait/instructional	paper-pencil	84
Thomas-Kilman Conflict MODE Survey	Thomas & Kilman (1974)	conflict management	college, adult	trait/instructional	paper-pencil	85
Various Dimensions of Communication						
Basic Course Communication Competency Measure	Ford & Wolvin (1993)	hybrid communication course	college	state/instructional	paper-pencil	86
Communicative Adaptability Scale	Duran (1983)	communicator style	college, adult	state/instructional	paper-pencil	87
The Communication Behaviors Inventory	Univ of Colorado at Colorado Springs, Center for Excellence in Oral Communication	overall communication behaviors	college	trait/instructional	paper-pencil and computerized self-scoring version available	88
Communication Competency Self-Report Questionnaire	Rubin (1985)	overall self-perception of communication competence	college	state/instructional	paper-pencil	89
Communication Flexibility Measure	Martin & Rubin (1994)	adapting communication behaviors	college, adult	trait/instructional	paper-pencil responses to behavioral scenarios	90
Focal Person's Communications Survey/Colleague Questionnaire	Klauss & Bass (1982)	communicator style	college, adult	trait/instructional	paper-pencil and other report by colleagues	91
I ACCESS: Assessment of Communication Competency & English Speaking Skills	Norma Landa Flores	various dimensions of communication	college, adults (ESL)	trait/instructional	videotaped subjects respond to visual/auditory prompts; raters score on a form	92
Norton Communicator Style Measure (CSM)	Norton (1978)	communicator style	college, adult	trait/instructional	paper-pencil	93
Openness Scale	Norton & Montgomery (1982)	willingness to share personal information	college	trait/instructional	paper-pencil	94
Self-Perceived Communication Competence Scale	McCroskey & McCroskey (1988)	overall perception of competence as a communicator	college, adult	state/instructional	paper-pencil	95
Source Credibility Scale	McCroskey (1994)	perceptions of source credibility	college, adult	trait/instructional	paper-pencil evaluations of a target	96
Willingness to Communicate Scale	McCroskey & Richmond (1987)	willingness to communicate in various contexts	college, adult	trait/instructional	paper-pencil	97

BEST COPY AVAILABLE

Kentucky Comprehensive Listening Test

Authors/Developers: R. N. Bostrom, & E. S. Waldhart. Ky. Listening Research Center, Univ of Kentucky, Lexington KY 40506. 606-257-7800.

Intended context or behavior: Listening

Intended academic level: College/Adult.

Time required: 60-90 minutes.

Text description: This instrument includes an audiotape and a four-part question/response sheet. The 42 items in Part I ask respondents to identify immediately a particular number or letter from a series heard on the audiotape. Similarly the 12 items in Part II ask respondents to identify a letter or number from a series but with a pause included. Part III includes four short interactions and asks respondents to complete 10 multiple-choice items about them. In Part IV respondents listen to a fairly extensive lecture and then respond to 14 multiple-choice questions.

Scoring description: Correct items are totaled to yield a global score. In addition, five subscales are available, which indicate the specific listening strengths and weaknesses of the respondents. The subscales measure short-term listening, short-term listening with rehearsal, interpretative listening, informative listening, and the ability to overcome distraction when listening. The materials provide means and percentile rankings for the global score and the subscales.

Norm/criterion reference populations: Forms of the scale have been administered to over 20,000 college students and adults. Subgroups were used for various statistical analyses.

Culture and gender bias: Not available.

Validity

Concurrent: The authors report moderate correlations between two of the parts (the short-term with rehearsal and the lecture section scores) and student performance on a speech course examination. Others have reported that this measure correlates with the Watson-Baker Listening Test and the listening items on the Communication Competency Assessment Instrument.

Face/content: University of Kentucky Communication Department faculty helped to validate the interaction section of the test. Although there was not unanimous agreement, these faculty determined the appropriateness of the response options.

Construct: Comparisons of the KCLT with other tests of cognitive ability suggest that it is measuring a distinct ability different from others that are closely related.

Reliability

Alternate forms: Reliability between Forms A and B was .72. The interpretative subscale (Part 3) however, for the alternative forms yielded .36.

Test/retest: Test-retest coefficients on an earlier version of this scale ranged from .78 to .87 for the various subscales.

Internal Consistency: Cronbach alpha coefficients for global scores on Forms A and B were .78 to .82. Subscales ranged from .46 to .76.

Evaluative Reactions

Practicality and usefulness: This measure is a very straightforward and economical way to assess the listening proficiency of a large number of people.

Clarity of instructions: The written form of the audiotape is clear and easy to follow. Directions for answering questions were clear.

Degree of standardization in administration and evaluation of responses: The test administrator hands out the test booklets and turns off the tape. It was not clear whether the questions and response choices were read aloud or just responded to in written form. Other researchers have reported a problem with respondents reading ahead.

Range of responses or degrees of competence: There is only one "right" answer to each question or item on the test. This seemed especially problematic in one section, where more than one response seemed appropriate.

Primary strengths: The stimulus requires that subjects base their answers on listening to the audio tape. This allows for control over the testing situation which a live presentation may not. The lecture and conversation situations are ones that are familiar to most students.

Primary weaknesses: The letter and number sequence sections are not naturalistic stimuli. Test scores may be complicated by reading ability. The lecture seemed long and uninteresting.

Overall adequacy: This is a viable option for departments who must conduct large scale assessment of listening proficiency.

Additional References:

Bostrom, R. N., & Waldhart, E. S. (1988). Memory models and the measurement of listening. *Communication Education, 37*, 1-13.

Clark, A. J. (1989). Communication confidence and listening competence: An investigation of the relationships of willingness to communicate, communication apprehension, and receiver apprehension to comprehension of content and emotional meaning in spoken messages. *Communication Education, 38*, 237-236.

Prather, M. M., Bostrom, R. N. (1991). *Birth order and listening: A preliminary investigation*. Lexington: University of Kentucky. (ERIC Document Reproduction Service No. ED 333 490)

Prather, M. M., Bostrom, R. N. (1991, April). *Birth order and listening: A preliminary investigation*. Paper presented at the meeting of the Southern States Communication Association, Tampa, FL.

Rubin, R. B., & Roberts, C. V. (1987). A comparative examination and analysis of three listening tests. *Communication Education, 36*(2), 142-153.

Rubin, R. B., & Shepherd, P. (1985). *Listening assessment via the Communication Competency Assessment Instrument*. Kent, OH: Kent State University. (ERIC Document Reproduction Service No. ED 264 616)

Rubin, R. B., & Shepherd, P. (1985, November). *Listening assessment via the Communication Competency Assessment Instrument*. Paper presented at the meeting of the Speech Communication Association, Denver, CO.

Reviewer: Ellen A. Hay.

BEST COPY AVAILABLE

Steinbrecher-Willmington Listening Test

Authors/Developers: M. M. Steinbrecher, & J. C. Willmington. Dept of Communication, Univ of Wisconsin, Oshkosh WI 54901. 414-424-7126.

Source: Information not available.

Cost: $189.50. Includes videotape, manual, and 30 rater sheets.

Intended context or behavior: Listening abilities.

Intended academic level: College.

Time required: 45 minutes.

Text description: This test includes a videotape and 55 multiple-choice questions divided into two parts. Part I asks students to watch and listen to a 4-minute speech, three announcements, directions, and a description. Part II contains 5 two-person conversations, and five statements containing examples of evidence and reasoning. After each segment, the tape is paused so that students can respond to the corresponding multiple-choice questions.

Scoring description: The number of correct responses is totaled. The test manual notes a mean score of 34.77 and also provides percentile ranks for scores.

Norm/criterion reference populations: Percentiles are based upon a sample of 916 college students. Tests of reliability and validity were performed with smaller samples of this group.

Culture and gender bias: The scores of men and women were not significantly different, although women consistently averaged one more correct answer.

Validity

Concurrent: Correlations ranging from .61 to .65 were found between the SWLT and the Watson-Baker Listening Test.

Face/content: Test developers relate the instrument to the various purposes of listening identified in the course text and other listening sources. Significant discussion also focuses on suggestions and findings on listening assessment.

Construct: Honors students tested significantly higher than regular students. Students receiving an A or B in a small-group class scored significantly higher than students receiving a D or F.

Reliability

Test/retest: With a 10-day interval, the test-retest coefficient was .68.

Internal consistency: A reliability coefficient of .67 was reported for the 916-member sample.

Evaluative Reactions

Practicality and usefulness: This measure considers different types of listening and would be a good overall measure of listening proficiency.

Clarity of instructions: Directions were clear.

Standardization in administration and evaluation: The one barrier to standardization may be the availability and applicability of the stimulus tape.

Range of responses or degrees of competence: Percentile ranks provide a range for assessing listening competence.

Primary strengths: Because a videotape is used, the assessment parallels "real" listening where both verbal and nonverbal information is available. The measure is based on a very comprehensive treatment of listening.

Primary weaknesses: This instrument has not been widely tested on populations other than students at the University of Wisconsin at Oshkosh.

Overall adequacy: This measure is a viable choice for departments that need to demonstrate or assess listening proficiency.

Additional References:

Steinbrecher, M. M. (1992, October). *Assessing listening proficiency*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Willmington, S. C., & Steinbrecher, M. M. (1993, April). *Assessing listening in the basic course: The University of Wisconsin-Oshkosh Listening Test*. Paper presented at the Joint Meeting of the Southern States Communication Association and the Central States Communication Association, Lexington, KY.

Reviewer: Ellen A. Hay.

Watson-Barker Listening Test

Authors/Developers: K. W. Watson, & L. L. Barker. Spectra, Inc., P. O. Box 5031, Contract Station 20, New Orleans LA 70118.

Source: Watson, K. W., Barker, L. L., & Roberts, C. V. (1989). *Development and administration of the High School Watson-Barker Listening Test: Facilitator's guide*. New Orleans: Spectra.

Cost: Adult versions A & B, $249.95; adult short version, $149.95.

Intended context or behavior: Listening.

Intended academic level: Adult.

Time required: 30 minutes for each form.

Text description: Each of two different forms (A and B) "includes 50 items equally divided into five sections: (1) evaluating message content; (2) understanding meaning in conversations; (3) understanding and remembering information in lectures; (4) evaluating emotional meanings in messages; and (5) following instructions and directions." Both forms are available in audio or video format. Preprinted test booklets are used with the audio form, and self-scoring worksheets only are used with the video form.

Scoring and administration description: Directions are provided for respondents to determine and interpret their own scores.

Norm/criterion reference populations: Included 20,000 undergraduates, primarily enrolled in speech classes, and 7,500 managers, supervisors, administrators, and professionals. Norms for the adult population are provided in the facilitator's manual.

Culture and gender bias: Females have slightly higher mean scores than males.

Validity

Predictive: Scores are correlated in expected ways with the Receiver Apprehension Test.

Face/content: Listening experts reviewed items.

Reliability

Alternate forms: Significant positive relationships are found between Versions A and B.

Evaluative Reactions

Practicality and usefulness: The instrument is easy to administer. It focuses on an important communication skill.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Directions are clear enough to lead to standardized usage.

Range of responses or degree of competence: Normative responses are available for comparison.

Primary strengths: Focuses on an understudied aspect of communication.

Primary weaknesses: Reliability and validity are not described in the facilitator's guide but may be obtained from the developers.

Additional References:

Karr, M., & Vogelsang, R. W. (1989). *A comparison of the audio and pilot video versions of the Watson-Barker Listening Test: Forms A & B*. Portland, OR: Portland State University. (ERIC Document Reproduction Service No. ED 304 729)

Karr, M., & Vogelsang, R. W. (1989, February). *A comparison of the audio and pilot video versions of the Watson-Barker Listening Test: Forms A & B*. Paper presented at the meeting of the Western States Speech Communication Association, Spokane.

Roach, D. A., & Fitch, H. M. (1984, March). *A comparison of three modes of administering listening tests*. Paper presented at the International Listening Association convention, Scottsdale, AZ.

Roberts, C. V. (1986, April). *The question of validity and application: How do we know how well we are listening*. Paper presented at the meeting of the Eastern Communication Association, Atlantic City.

Rubin, R. B., & Roberts, C. V. (1987). A comparative examination and analysis of three listening tests. *Communication Education, 36*, 142-153.

Rubin, R. B., & Shepherd, P. (1985). *Listening assessment via the Communication Competency Assessment Instrument*. Kent, OH: Kent State University. (ERIC Document Reproduction Service No. ED 264 616)

Rubin, R. B., & Shepherd, P. (1985, November). *Listening assessment via the Communication Competency Assessment Instrument*. Paper presented at the meeting of the Speech Communication Association, Denver.

Smeltzer, L. R., & Watson, K. W. (1984). Listening: An empirical comparison of discussion length and level of incentive. *Central States Speech Journal, 35*, 166-170.

Watson, K. W. (1984). Listening: Who's testing what? *Curriculum Review, 23*, 20-21.

Watson, K. W., & Barker, L. L. (1985, November). *Listening assessment: The Watson Barker Listening Test*. Paper presented at the annual meeting of the Speech Communication Association, Denver.

Watson, K. W., Barker, L. L., & Roberts, C. V. (1989). Watson Barker Listening Test. In J. C. Conoley & J. J. Kramer (Eds.), *The tenth mental measurements yearbook* (10th ed., pp. 384-385). Lincoln: Buros Institute of Mental Measurements, University of Nebraska-Lincoln.

Watt, W. M. (1993). *Theoretical and practical approaches to teaching listening: Using the Watson-Barker Listening Test to validate levels of learning in the classroom*. Fort Hays, KS: Fort Hays State Univ. (ERIC Document Reproduction Service No. ED 366 026)

Watt, W. M. (1993, November). *Theoretical and practical approaches to teaching listening: Using the Watson-Barker Listening Test to validate levels of learning in the classroom*. Paper presented at the meeting of the Speech Communication Association, Miami Beach.

Reviewer: Judith M. Dallinger.

BEST COPY AVAILABLE

Conversational Skills Rating Scale

Author/Developer: B. H. Spitzberg. School of Communication, San Diego State Univ, San Diego CA 92182.

Source: Speech Communication Association, 5105 Backlick Rd #F, Annandale VA 22003.

Cost: $16.50 for SCA members, $19 for nonmembers.

Intended context or behavior: Interpersonal communication skills.

Intended academic level: College/Adult.

Time required: 5-15 minutes for administration of the scale; additional time would be required if subjects are responding to an interaction stimulus prior to completing scale.

Text description: The scale consists of 30 items; 25 items focus on specific behaviors subdivided into four clusters of altercentrism, composure, expressiveness, and interaction management. Five items consider general impressions of the subject's performance. The scale can be completed after subjects engage in a stimulus conversational task or after subjects recall an important conversation. Friends and relatives can also complete the measure in reference to the subject. The material includes variations of the scale so that subjects can rate themselves, or they can be evaluated by a partner, a trained observer, or someone who is familiar with their conversational manner.

Scoring description: The 25 specific behaviors are rated on a 1-5 scale: 1 = inadequate, 2 = fair, 3 = adequate, 4= good, 5 = excellent. Five general items rate performance on a 7-point scale. The five general items can be used to validate the specific items. Cutoff points for competent versus incompetent performance are not included. The author suggests that, rather than as a general label, the scale is better used to diagnosis discrete communication strengths and weaknesses. Scores can also be compiled as an indication of program effects. Four subscales can also be used for instructional purposes.

Norm/criterion reference populations: "The CSRS has been applied in groups of students, business persons, nonstudent adults, and military personnel. It has been applied to problem-solving interactions, laboratory situations, naturally-occurring conversations, and get acquainted conversations. It has been applied as a state, a cross-contextual traits, and a relationship-based disposition" (p. 11).

Culture and gender bias: Specific tests were not highlighted in materials accompanying the scale. Studies on an earlier version reported that females perceived themselves more competent on several items in the measure. Female raters tended to rate subjects higher than male raters.

Validity

Concurrent: The author reports that studies have shown significant correlations between the CSRS and measures of conversational satisfaction, relational trust, received social support, use of humor, loneliness, trait apprehension, and partner perceptions of actor competence.

Face/content: Items are selected by an extensive four-stage process discussed in the materials.

Construct: The author reports that the CSRS "has generally produced validity coefficients in the expected direction and or reasonable magnitude" (SCA, p. 10). The material includes an appendix that presents results from 21 studies of the instrument at various stages of its development.

Reliability

Test/retest: : A .78 correlation resulted when given on a pretest-posttest basis.

Scoring: Acceptable interrater reliabilities have been reported.

Internal consistency: Consistency was consistently over .80.

Evaluative Reactions

Practicality and usefulness: The scale is an efficient way of measuring interpersonal skills so that areas needing improvement can receive attention. It would also be useful as one indicator of the effectiveness of a course or program.

Clarity of instructions: Rating procedures and scale items are clear and well-defined.

Standardization in administration and evaluation: The material includes a training guide that suggests training procedures and elaborates on each of the 25 specific behaviors in terms of the five possible ratings.

Range of responses or degrees of competence: Labeling of performance is not encouraged.

Primary strengths: This is a very user-friendly measure. It is clear and easily adapted to a number of uses. It is based upon a rational definition of interpersonal competence and acknowledges the role of context and culture in competence.

Primary weaknesses: More attention to cultural and gender differences would be helpful.

Additional References

The materials from SCA contain an extensive bibliography.

Canary, D. J., & Spitzberg, B. H. (1987). Appropriateness and effectiveness perceptions of conflict strategies. *Human Communication Research, 14*, 93-118.

Canary, D. J., & Spitzberg, B. H. (1989). A model of the perceived competence of conflict strategies. *Human Communication Research, 15*(4), 630-649.

Canary, D. J., & Spitzberg, B. H. (1990). Attribution biases and associations between conflict strategies and competence outcomes. *Communication Monographs, 57*(2), 139-151.

Ellis, D. G., Duran, R. L., & Kelly, L. (1994). Discourse strategies of competent communicators: Selected cohesive and linguistic devices. *Research on Language and Social Interaction, 27*, 145-170.

Spitzberg, B. H., & Hecht, M. L. (1984). A component model of relational competence. *Human Communication Research, 10*, 575-599.

Spitzberg, B. H., & Hurt, H. T. (1987). The measurement of interpersonal skills in instructional contexts. *Communication Education, 36*, 28-45.

Reviewer: Ellen A. Hay.

The Competent Speaker

Authors/Developers: S. P. Morreale, M. R. Moore, K. P. Taylor, D. S. Tatum, & R. Hulbert-Johnson. Speech Communication Association, 5105 Backlick Rd #F, Annandale VA 22003. 703-750-0533.

Source: Morreale, S. P., Moore, M. R., Taylor, K. P., Surges-Tatum, D., & Hulbert-Johnson, R. (1993). *The Competent Speaker Speech Evaluation Form*. Annandale, Va: Speech Communication Association.

Cost: $22.50 for nonmembers, $17.50 for SCA members. Must be prepaid.

Intended context or behavior: Public Speaking.

Intended academic level: College.

Time required: Training requires approximately 2 hours. Administration of the instrument requires the length of the assigned speech plus an additional 10 minutes.

Text description: This instrument presents a standardized form and approach for evaluating public speeches. It is based upon eight competencies that include: topic, purpose, supporting material, organization, language, voice, usage, and physical behaviors. The training manual includes a discussion of each competency and an explanation of how each would be demonstrated at excellent, satisfactory, and unsatisfactory levels. Instructions are included for preparing a videotape to demonstrate excellent, satisfactory, and unsatisfactory speeches, along with information on how other speech communication educators evaluated sample speeches.

Scoring description: After receiving training, an evaluator rates student speeches on each of the eight competencies. Points are assigned. A rating of unsatisfactory receives one point, satisfactory receives two points, and excellent receives three points. Scores range from 8-24. Users can set their own standards for determining competency. As explained in the training manual, scoring can also be adjusted to facilitate attention to particular competencies.

Norm/criterion reference populations: Twelve speech communication educators, 10 graduate teaching assistants, and three community college instructors provided referencing scores for the three anchor speeches along with the other sample speeches on the training tape.

Culture and gender bias: Two studies were conducted to investigate bias. In one study, minority students rated the sample speeches. No differences were found between their ratings and the ones reported by the speech communication educators. When the instrument was used in a speech class with 260 students, no significant differences in scores were found among students from different cultural groups. No significant differences were reported between men and women.

Validity

Concurrent: A negative correlation was reported between this instrument and the six public speaking items on the Personal Report of Communication Apprehension, indicating that better public speakers had lower apprehension. A positive correlation was reported with the seven public speaking items on the Communication Competency Assessment Instrument.

Face/content: Developers conducted an extensive literature review when determining appropriate competencies and criteria. A panel of 11 speech communication educators was involved in the final determination.

Reliability

Scoring: High interrater reliability was reported after training of evaluators.

Evaluative Reactions

Practicality and usefulness: The developers suggest many possible uses for this instrument. Using this form to evaluate all student speeches and maintaining records of this use will be invaluable for departments that must demonstrate that students are learning to speak in public.

Clarity of instructions: Competencies are clear as are rating procedures. Description of Rasch Analysis is not entirely clear but not really necessary to the use of the instrument.

Standardization in administration and evaluation: Content of the training session might be a little more definite. Each competency is fully explained to allow for optimal standardization when different evaluators use the instrument.

Range of responses or degrees of competence: Assigning the label of "competent" is done on a scale.

Primary strengths: This is a clear, simple way to bring uniformity to speech evaluations.

Primary weaknesses: The instrument may not highlight thinking, reasoning, and evidence as substantial components of speech preparation and presentation. As we focus more on the connection between critical thinking and communication, this link may need to be strengthened. More space for written feedback is also needed.

Overall adequacy: The Competent Speaker will be a valuable tool for many departments. It could be particularly effective for institutions that offer multiple sections of the public speaking course because it allows for standardized evaluation of the speeches.

Additional References

Hugenberg, L. W., & Yoder, D. D. (1994). *Communication competence: A reaction to the "Competent Speaker Speech Evaluation Form"*. Youngstown, OH: Youngstown State University. (ERIC Document Reproduction Service No. ED 372 432)

Hugenberg, L. W., & Yoder, D. D. (1994, April). *Communication competence: A reaction to the "Competent Speaker Speech Evaluation Form"*. Paper presented at the meeting of the Eastern Communication Association, Washington, DC.

Morreale, S. (1990). *"The competent speaker": Development of a communication-competency based speech evaluation form and manual*. Colorado Springs: University of Colorado at Colorado Springs. (ERIC Document Reproduction Service No. ED 325 901)

Morreale, S. (1990, November). *"The competent speaker": Development of a communication-competency based speech evaluation form and manual*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Morreale, S. P., Awtrey, C. C., Hulbert-Johnson, R., Morley, D. D., Moore, M. R., Surges-Tatum, D., & Taylor, K. P. (1991). *The competent speaker: A workshop on assessing public speaking competence*. Colorado Springs: University of Colorado at Colorado Springs. (ERIC Document Reproduction Service No. ED 344 253)

Morreale, S. P., Awtrey, C. C., Hulbert-Johnson, R., Morley, D. D., Moore, M. R., Surges-Tatum, D., & Taylor, K. P. (1991, November). *The competent speaker: A workshop on assessing public speaking competence*. Short course presented at the meeting of the Speech Communication Association, Atlanta.

Morreale, S. P., Awtrey, C. C., Hulbert-Johnson, R., Morley, D. D., Moore, M. R., Surges-Tatum, D., & Taylor, K. P. (1993, November). *The competent speaker: A workshop on assessing public speaking competence*. Short course presented at the meeting of the Speech Communication Association, Miami.

Morreale, S. P., Gomez, A., Shockley-Zalabak, P., Lowe, H., & Pinello-Tegtmeier, L. (1992). *Oral communication competency assessment: Cultural diversity, ethnicity, and students at-risk*. Colorado Springs: University of Colorado at Colorado Springs. (ERIC Document Reproduction Service No. ED 350 642)

Morreale, S. P., Gomez, A., Shockley-Zalabak, P., Lowe, H., & Pinello-Tegtmeier, L. (1992, November). *Oral communication competency assessment: Cultural diversity, ethnicity, and students at-risk*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Morreale, S. P., Whitney, P., Zautke, B., Ellis, K., McCormick, C., & Witter, S. (1992). *The Center for Excellence in Oral Communication: A comprehensive program for assessing the development of public speaking competency*. Colorado Springs: University of Colorado at Colorado Springs. (ERIC Document Reproduction Service No. ED 350 642)

Morreale, S. P., Whitney, P., Zautke, B., Ellis, K., McCormick, C., & Witter, S. (1992, November). *The Center for Excellence in Oral Communication: A comprehensive program for assessing the development of public speaking competency*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Reviewer: Ellen A. Hay.

Dantes Principles of Public Speaking

Authors/Developers: Educational Testing Service. The Dantes Program, Educational Testing Services, Mail Stop 23-N, Princeton NJ 08541-0001. 609-951-6264.

Source: Information not available.

Cost: $10 per student tested.

Intended context or behavior: Public speaking examination used to award students course credit for equivalent knowledge.

Intended academic level: College/Adult.

Time required: 90-120 minutes.

Text description: This is a two-part examination. Part I includes 84 multiple-choice items that question respondent knowledge on various aspects of public speaking: ethical, historical and social considerations, audience analysis and adaptation, topics and purposes, organization, supporting material, research, language and style, and delivery. Part II requires that same respondent to deliver, after 10 minutes of preparation, a 3-5 minute impromptu speech (the speech is audiotaped).

Scoring description: The first part is sent to ETS for scoring. A minimum score of 49 is recommended for awarding three college credits in public speaking. Schools may set their own standards for the awarding of credits. Part II is graded on a pass/fail basis using the criteria of persuasive strategies, topic focus, organization, language, oral delivery, and length. It can be evaluated through ETS or by the institution awarding credit.

Norm/criterion reference populations: College students (N = 1,013) from 55 public and private colleges and universities who had recently completed an elementary college public speaking course were used as the norm referencing group.

Culture and gender bias: An insufficient number of examinees in the reference group were of varying ethnic backgrounds so no analysis of cultural bias was conducted. There was a ½ point difference reported in the scores of men and women.

Validity

Face/content: Questions were based on material covered in the commonly used public speaking texts and were reviewed by a panel of speech communication educators.

Construct: A positive correlation (.44) was reported between test scores and grades of the norm referencing population.

Reliability

Internal consistency: A reliability coefficient of .87 was reported.

Evaluative Reactions

Practicality and usefulness: This is a highly standardized approach that would be helpful to institutions who want to award credit for experiential learning or who want to validate their own public speaking course outcomes.

Clarity of instructions: Instructions are very clear.

Standardization in administration and evaluation: Directions and evaluation procedures are highly standardized.

Range of responses or degrees of competence: The American Council on Education has established a standard score of 47 as passing. Institutions may establish their own levels.

Primary strengths: This is a very thorough approach to assessing public speaking ability. It includes both the knowledge and ability necessary in speaking.

Primary weaknesses: The speech is audiotaped so many physical qualities are not considered in the evaluation process.

Overall adequacy: Because of the significant research which has gone into the development of this instrument, it provides a very thorough evaluation of public speaking knowledge.

Additional References

Jones, D. W., & Saltman, L. E. (1989). Educating the military work force: A worldwide initiative. *Adult Learning, 1*, 20-21.

Office of the Assistant Secretary of Defense for Force Management and Personal, Education Programs in the Department of Defense. (1990). Washington, DC: GPO.

Reviewer: Ellen A. Hay.

Speaking Skills Assessment College Outcome Measures Program

Authors/Developers: American College Testing, P. O. Box 168, Iowa City IA 52243.

Source: Information not available.

Cost: $3 to $6 per student, depending on the number ordered.

Intended context or behavior: Public speaking role-play to large and small audiences.

Intended academic level: College.

Time required: 30 minutes is needed for training. Each speaker evaluation takes 15-30 minutes.

Text description: Respondents are asked to prepare three short informative or persuasive speeches appropriate to stimulus situations given to the respondent one day in advance. Respondents then return to the testing site where the presentations are audiotaped.

Scoring description: Raters either from the institution or from ACT (at a cost of $5 per respondent) evaluate each audiotaped presentation on three criteria: audience, discourse, and delivery. Each criterion has five levels of proficiency which are described in materials provided by ACT. Based upon assigned ratings, each respondent then receives from ACT a total score along with subscores for each criterion.

Norm/criterion reference populations: Initially, 1,589 first-year students and 1,366 graduating students were involved in norm referencing scores.

Culture and gender bias: One validation study included data on "effectively functioning adults" from diverse cultural groups. Correlations between the measure and the effectiveness index were consistent despite cultural differences. The rating scale seems to offer enough definition to discourage rating bias. Pronunciation and dialect were not mentioned in delivery criteria. Validation studies indicate that women perform better than do men.

Validity

Face/content: The measure was designed with the intention of making it "authentic."

Construct: Six studies were conducted to demonstrate construct validity. These studies were intended to show that adults who functioned effectively in a variety of contexts possessed the abilities measured by this assessment. Moderate correlations were reported in these studies. When used as a pre- and postcourse measure, the Speaking Skills Assessment found that students' abilities improved significantly after completing a rhetoric course which included study and practice in writing and speaking. Improvement was noticeable for those whose precourse scores were low. Differences in scores have also been reported for 1st-year and graduating students, and for students with various grade point averages and ACT scores.

Reliability

Alternate forms: Correlation coefficients ranging from .71 to .81 were reported for six different forms of the measure.

Scoring: Interrater reliability of .92 was reported for a group of 15 evaluators.

Internal consistency: Estimates of internal consistency were high.

Evaluative Reactions

Practicality and usefulness: This instrument provides a standardized approach for evaluating student speaking. Because of the vast study and subsequent norm referencing of this instrument, it is particularly beneficial to departments who want some basis of comparison between their students and other students from around the country.

Clarity of instructions: Instructions in the simulations seem very clear.

Standardization in administration and evaluation: The criteria for rating responses were a bit vague at some points, in such statements as "making contact with the relevant audience" or "structures a balanced discourse.".

Range of responses or degrees of competence: Each criterion is evaluated at five levels of proficiency.

Primary strengths: This instrument has undergone extensive testing and study. Alternate forms are available and rater responses can be checked for consistency.

Primary weaknesses: Because only audiotapes are evaluated, few of the nonverbal elements of speaking can be considered. The instrument might produce undue stress. Highly apprehensive students probably will not show up on presentation day after reading the simulation instructions.

Overall adequacy: This is a very viable option for assessing public speaking abilities.

Additional References

College Outcome Measures Program. (1992). *Technical report 1982-1991: Clarifying and assessing general education outcomes of college*. Iowa City: American College Testing Program.

Trank, D., & Steele, J. M. (1983). Measurable effects of a communication skills course: An initial study. *Communication Education, 32*, 227-236.

Reviewer: Ellen A. Hay.

Behavioral Assessment of Speech Anxiety (BASA)

Authors/Developers: A. Mulac, & R. Sherman. Anthony Mulac, Communication Dept, Univ of California, Santa Barbara CA 93106.

Source: Mulac, A., & Sherman, R. (1974). Behavioral assessment of speech anxiety. *Quarterly Journal of Speech, 60*, 134-143.

Cost: Free/public domain.

Intended context or behavior: Communication apprehension.

Intended academic level: College/Adult.

Time required: In the validating study, raters trained with a 58-minute videotape of speeches. Communication professionals should be able to use the instructions included in the description article.

Text description: A rating form which includes 18 behavioral scales on which observers rate speakers (from videotapes). Behaviors are observed and rated on 9-point scales for 1-minute segments of speeches.

Scoring description: Factor analysis produced four different dimensions which can be summed separately. Scales are weighted as described in the source study.

Culture and gender bias: Information not available.

Validity

Predictive: Speeches by 10 males were rated as less anxious on the fourth speech they presented in a public speaking class, as compared to their first speeches.

Concurrent: Total BASA scores correlated with speech anxiety scores from 18 validation judges at .88.

Reliability

Scoring: Weighted total BASA scores were .95. Interrater reliability for each behavioral variable ranged from .70 to .96, except for one variable which was .35. Average interrater reliability across behaviors was .88.

Evaluative Reactions

Practicality and usefulness: The rating scales are difficult to use easily. They do provide more precise measures of total anxiety scores than an individual scale.

Clarity of instructions: The directions are clear for speech raters.

Standardization in administration and evaluation: Ratings were similar for 18 different judges.

Primary strengths: It measures specific behavioral aspects of speech anxiety.

Primary weaknesses: It has been used seldom as a research tool, never as an assessment technique.

Overall adequacy: This has potential to be developed as an assessment technique.

Additional References

Romero, G. (1982). *An experimental study of the effects of counterattitudinal messages on decreasing communication apprehension*. Reno: University of Nevada. (ERIC Document Reproduction Service No. ED 221 897)

Romero, G. (1982, November). *An experimental study of the effects of counterattitudinal messages on decreasing communication apprehension*. Paper presented at the meeting of the Speech Communication Association, Louisville.

Sherman, A. R., Mulac, A., & McCann, M. J. (1974). Synergistic effect of self-relaxation and rehearsal feedback in the treatment of subjective and behavioral dimensions of speech anxiety. *Journal of Consultation and Clinical Psychology, 42*(6), 819-827.

Reviewer: Judith M. Dallinger.

Communication Competency Assessment Instrument Revised Edition

Author/Developer: R. B. Rubin, School of Communication Studies, Kent State Univ, P. O. Box 5190, Kent OH 44242-0001; or Spectra, Inc., 4197 Monterey Drive, Memphis TN 38128. 901-386-7666.

Source: Information not available.

Cost: Starter pack = $185 (includes 2 manuals, 50 rating sheets, and 1 videotape. Individual items: 1 package of rating sheets = $9.50; 1 videotape = $85; 1 manual = $15.

Intended context or behavior: Public Speaking, interaction abilities, and listening.

Intended academic level: College.

Time required: 30-45 minutes.

Text description: The CCAI includes three parts. Subjects are given 5 minutes to prepare a 3-minute speech on a topic of their choice, which is then presented in a classroom setting to a rater. Students then view a videotape that simulates the first day of a fictitious college class. The rater discusses the tape with the student, asking specific questions to probe listening comprehension and interaction abilities.

Scoring description: The raters assess 19 competencies during the speech and interview sessions. The subject is evaluated on a 5-point scale for each competency. For each of the 19 competencies, raters are provided with clear descriptions of each of the five levels. Rubin has cautioned that the instrument should be considered as a global measure of communication competence, although individual items may indicate strengths and weaknesses.

Norm/criterion reference populations: Various groups of college students have been used in the development and testing of this instrument.

Culture and gender bias: Validation data indicate that the test does not appear to be biased on the basis of gender or ethnic differences. In the technical manual, Rubin discusses many of the common forms of bias that should be addressed in rater recruiting, training, and monitoring.

Validity

Predictive: Previous studies have found that the CCAI was an effective predictor of success for student teachers and have reported that those with higher CCAI scores were more likely to persist in college. The CCAI has been positively correlated with GPA, high-school speaking experiences, and college speaking experiences when the subjects were freshmen and sophomores.

Concurrent: Moderate correlations have been reported between the CCAI and the Communication Competence Self-Report Questionnaire, although most part subjects seem to overestimate their competence.

Face/content: The domain of 19 competencies was drawn from previous taxonomies of functional communication abilities. Speech communication experts considered whether the specific items on the assessment accurately tapped the abilities they were intended to measure.

Reliability

Scoring: Interrater reliability coefficients have been good. The rating scale contains good descriptions of the variations of behaviors being assessed, and consistency was high among raters, particularly if they receive training.

Evaluative Reactions

Practicality and usefulness: This is one of the few measures available which actually evaluates the communication abilities of the subject, rather than relying on a self-report of these abilities. The drawback is that a tremendous time commitment is necessary when using this measure on a large scale.

Clarity of instructions: The directions and probes are very clear. Further research has found that adding a bit more information to each probe can improve CCAI performance.

Standardization in administration and evaluation: The descriptions of each level of each competency are very definite and provide for a standardized approach.

Range of responses or degrees of competence: The evaluation scale allows for a variety of responses rather than just one right answer.

Primary strengths: Because the probes are spoken and the responses are oral, the measure does not test reading and writing abilities. The situations used in the probes are familiar to students, and the classroom setting is naturalistic. The stress of the testing situation can be kept to a minimum, especially if the rater is effectively trained. The stimulus for the listening test is videotaped, and the responses to this section are minimal and do not test memory.

Primary weaknesses: Presently there are no alternative forms of this measure. The public speaking portion is not given to an audience and therefore might not be a totally accurate assessment of a student's speaking ability.

Overall adequacy: The measure would be especially effective for departments seeking to assess the communication abilities for a representative sample of students.

Additional References

Carpenter, E. C. (1986). *Measuring speech communication skills.* (ERIC Document Reproduction Service No. 268 599)

Carpenter, E. C. (1986, April). *Measuring speech communication skills.* Paper presented at the meeting of the Central States Speech Association, Cincinnati.

Powell, R. G., & Avila, D. R. (1986). Ethnicity, communication competency and classroom success: A question of assessment. *Western Journal of Speech Communication, 50,* 269-278.

Rubin, R. B. (1981). *The development and refinement of a communication competency assessment instrument.* Kent, OH: Kent State University. (ERIC Document Reproduction Service No. 210 732)

Rubin, R. B. (1981, November). *The development and refinement of a communication competency assessment instrument.* Paper presented at the meeting of the Speech Communication Association, Anaheim, CA.

Rubin, R. B. (1985a). Communication assessment instruments and procedures in higher education. *Communication Education, 33*(2), 178-180.

Rubin, R. B. (1985b). The validity of the communication competency assessment instrument. *Communication Monographs, 52,* 173-185.

Rubin, R. B. (1986). A response to "Ethnicity communication competency and classroom success. *Western Journal of Speech Communication, 50*(3), 279-282.

Rubin, R. B., & Graham, E. E. (1986). *Communication determinants of college success: An exploratory investigation.* Kent, OH: Kent State University. (ERIC Document Reproduction Service No. 278 073)

Rubin, R. B., & Graham, E. E. (1986, November). *Communication determinants of college success: An exploratory investigation.* Paper presented at the meeting of the speech Communication Association, Chicago.

Rubin, R. B., & Graham, E. E. (1988). Communication correlates of college success: An exploratory investigation. *Communication Education, 37*(1), 14-27.

Rubin, R. B., Graham, E. E., & Mignerey, J. T. (1990). A longitudinal study of college students' communication competence. *Communication Education, 39,* 1-14.

Rubin, R. B., & Henzl, S. A. (1984). Cognitive complexity, communication competence, and verbal ability. *Communication Quarterly, 32*(4), 263-270.

Rubin, R. B., & Roberts, C. V. (1987). A comparative examination and analysis of three listening tests. *Communication Education, 36*(2), 142-153.

Rubin, R. B., & Shepherd, P. (1985). *Listening assessment via the Communication Competency Assessment Instrument.* Kent, OH: Kent State University. (ERIC Document Reproduction Service No. 264 616)

Rubin, R. B., & Shepherd, P. (1985, November). *Listening assessment via the Communication Competency Assessment Instrument.* Paper presented at the meeting of the Speech Communication Association, Denver.

Trank, D. M. (1988). *Standardized tests as a measure of value-added to communication students.* Iowa City: University of Iowa. (ERIC Document Reproduction Service No. 299 631)

Trank, D. M. (1988, November). *Standardized tests as a measure of value-added to communication students.* Paper presented at the meeting of the Speech Communication Association, New Orleans.

Reviewer: Ellen A. Hay.

Listener Preference Profile

Authors/Developers: K. W. Watson, L. L. Barker, & J. Weaver. Spectra, Inc., P. O. Box 5031, Contract Station 20, New Orleans LA 70118; or Pfeiffer & Co., San Diego CA 92138.

Source: Watson, K. W., & Barker, L. L. (1993). *Guide to using the listener preference profile: Tips for trainers and facilitators.* New Orleans: Spectra, Inc.

Cost: Instruments: 10/$39.95, 100/$299.95; Facilitator's Guide: $24.95.

Intended context or behavior: Listening Style.

Intended academic level: Adult version.

Time required: 15-20 minutes to complete the instrument and figure one's own score.

Text description: Twenty items measure four self-reported preferences of listening including: people oriented, content oriented, action oriented, and time oriented. Results for a given person may indicate no listening preference (listening avoidance), one listening preference, or multiple listening preferences.

Scoring and administration description: Directions are provided for respondents to determine and interpret their own scores.

Norm/criterion reference populations: Forty percent of respondents have two or more listening preferences. Thousands of participants have been field tested during interactive listening seminars and other programs throughout the U.S.

Culture and gender bias: Gender differences have been documented in listening preferences. Women are more often people-oriented, while men are more often content- or action-oriented. More men than women have no listener preference (listening avoidance).

Validity

Face/content: Items were selected based on systematic literature review, and were reviewed by experts. .

Reliability

Information not available.

Evaluative Reactions

Practicality and usefulness: The scores on the instrument promote more careful thinking about one's own and other's listening motivations.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Directions for administering and evaluating are clear enough to lead to standardized usage.

Range of responses or degree of competence: Not much is known about normative responses.

Primary strengths: Focuses on an understudied aspect of communication.

Primary weaknesses: Reliability and validity information is provided in unpublished work.

Additional References

Watson, K. W., Watson, K. W., Barker, L. L., & Weaver, J. B., III. (1992). *Development and validation of the Listener Preference Profile.* New Orleans: Tulane University. (ERIC Document Reproduction Service No. ED 348 703)

Watson, K. W., Watson, K. W., Barker, L. L., & Weaver, J. B., III. (1992, March). *Development and validation of the Listener Preference Profile.* Paper presented at the meeting of the International Listening Association, Seattle.

Reviewer: Judith M. Dallinger.

Affective Communication Test

Authors/Developers: H. S. Friedman, L. M. Prince, R. E. Riggio, & M. R. DiMatteo. Howard S. Friedman, Dept of Psychology, Univ of California, Riverside CA 92521.

Source: Friedman, H. S., Prince, L. M., Riggio, R. E., & DiMatteo, M. R. (1980). Understanding and assessing nonverbal expressiveness: The affective communication test. *Journal of Personality and Social Psychology, 39*, 333-351.

Cost: Information not available.

Intended context or behavior: Nonverbal expressiveness during interpersonal interaction.

Intended academic level: College/Adult.

Time required: 5-10 minutes.

Text description: This self-report measure includes 13 items. Respondents indicate on a 9-point Likert-type scale the degree to which each statement applies to them.

Scoring description: Scoring for six of the items is reversed and then a total score is computed. Scores ranged from 28-114.

Norm/criterion reference populations: Various groups of college students were used in developing and testing this measure.

Culture and gender bias: Women tend to score slightly higher than men.

Validity

Predictive: ACT scores highly correlated with perceptions by others of the individual's expressiveness. High ACT scores were correlated with individuals who frequently engage in behavior intended to "move, inspire, and captivate others" (Friedman, Prince, Riggio, & DiMatteo, 1980, p. 337).

Face/content: The source article discusses conceptualization of the nonverbal expressiveness construct and past research in this area.

Construct: ACT correlated in the predicted manner with other measures of personality such as extroversion, neuroticism, social desirability, anxiety, self-esteem, and self-monitoring.

Reliability

Test/retest: After a 2-week interval, a test-retest correlation of .90 was reported.

Internal consistency: A reliability coefficient of .77 was reported.

Evaluative Reactions

Practicality and usefulness: This might be useful on a pre/post course basis as an indicator of behavioral change.

Clarity of instructions: Directions are very clear.

Standardization in administration and evaluation: The measure is highly standardized.

Range of responses or degrees of competence: Interpretative ranges are not provided in the article.

Primary strengths: This measure has undergone validation studies, reported in the source article, and offers a unique focus on the nonverbal dimension of interpersonal interaction.

Primary weaknesses: Users need to avoid overgeneralizing from this measure because it focuses on a fairly specific construct.

Overall adequacy: When used in conjunction with other measures, the ACT could be a good indicator of interpersonal effectiveness.

Additional References

Carpenter, E. C. (1986). *Measuring speech communication skills*. (ERIC Document Reproduction Service No. 268 599)

Carpenter, E. C. (1986, April). *Measuring speech communication skills*. Paper presented at the meeting of the Central States Speech Association, Cincinnati.

Friedman, H. S. (1991). Nonverbal display of emotion in public and in private: Self-monitoring, personality, and expressive cues. *Journal of Personality & Social Psychology, 61*(5), 766-775.

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality & Social Psychology, 51*(3), 649-660.

Riggio, R. E. (1988). Captivating, charming, cajoling...Charisma! *New Woman, 18*(5), 57-58.

Riggio, R. E., & Friedman, H. S. (1986). Impression formation: The role of expressive behavior. *Journal of Personality & Social Psychology, 50*(2), 421-427.

Trank, D. M. (1988). *Standardized tests as a measure of value-added to communication students*. Iowa City: University of Iowa. (ERIC Document Reproduction Service No. 299 631)

Trank, D. M. (1988, November). *Standardized tests as a measure of value-added to communication students*. Paper presented at the meeting of the Speech Communication Association, New Orleans.

Reviewer: Ellen A. Hay.

The Argumentativeness Scale

Authors/Developers: D. A. Infante, & A. S. Rancer. D. Infante, School of Communication Studies, Kent State Univ, Kent OH 42242.

Source: Infante, D. A., & Rancer, A. S. (1982). A conceptualization and measure of argumentativeness. *Journal of Personality Assessment, 46,* 72-80.

Cost: Free/public domain.

Intended context or behavior: Measures the tendency in individuals to approach or avoid arguments.

Intended academic level: College.

Time required: 15 minutes.

Text description: Respondents react to 20 statements on a 5-point Likert-type scale noting how well that statement describes their attitudes and behaviors in argumentative situations.

Scoring description: Two subscales are calculated: 10 summed items score the tendency to approach argumentative situations, 10 other items score the tendency to avoid arguments. The argumentative trait is calculated from the difference. A mean score of 4.44 was reported for the generalized argumentative trait. No ranges for high, low, and normal argumentativeness are given.

Norm/criterion reference populations: Subjects were college/university students.

Culture and gender bias: Information not available.

Validity

Predictive: High and moderate correlations were found between subject responses and the responses of a friend as to that subject's predisposition to argue.

Concurrent: Argumentativeness as measured by this scale was correlated slightly or moderately with measures of apprehension, willingness to communicate, and predispositions of verbal behavior.

Face/content: Research and theory contributing to scale development is described.

Construct: Subjects high in argumentativeness were found to be more likely to engage in argumentative situations.

Reliability

Test/retest: An $r = .91$ between test and retest scores indicates high stability.

Internal consistency: Cronbach alpha of .91 was reported for approach items and.86 for avoid items.

Evaluative Reactions

Practicality and usefulness: This instrument measures a fairly narrow predisposition to engage in argument, but it might be useful in assessing the success of instruction in argumentation, persuasion, and assertiveness.

Clarity of instructions: Very clear instructions are provided.

Standardization in administration and evaluation: Scoring of the instrument is straightforward and consistent.

Range of responses or degrees of competence: No ranges or judgments are provided for this scale.

Primary strengths: This is a straightforward measure that might be used as a part of in-class instructional assessment processes.

Primary weaknesses: No ranges are provided that would be useful in interpreting scores.

Additional References

The instrument is also available at ComServe.

Infante, D. A. (1989a). Argumentativeness and affirming communicator style as predictors of satisfaction/dissatisfaction with subordinates. *Communication Quarterly, 37*(2), 81-90.

Infante, D. A. (1989b). Response to high argumentatives: Message and sex differences. *Southern Communication Journal, 54*(2), 159-170.

Infante, D. A. (1992a). Initiating and reciprocating verbal aggression: Effects on credibility and credited valid arguments. *Communication Studies, 43*(3), 182-190.

Infante, D. A. (1992b). Verbal aggressiveness: Messages and reason. *Communication Quarterly, 40*(2), 116-126.

Infante, D. A. (1993). Relations between argumentative motivation, and advocacy and refutation on controversial issues. *Communication Quarterly, 41*(4), 415-426.

Infante, D. A. (1994). Argument and verbal aggression in constructive and destructive family and organizational disagreements. *Western Journal of Communication, 58*(2), 73-84.

Rancer, A. S. (1992). Beliefs about arguing as predictors of trait argumentativeness: Implications for training in argument and conflict management. *Communication Education, 41*(4), 375-387.

Sabourin, T. C. (1993). Verbal aggression in marriages: A comparison of violent, distressed but nonviolent, and nondistressed couples. *Human Communication Research, 20*(2), 245-267.

Suzuki, S., & Rancer, A. S. (1994). Argumentativeness 7 verbal aggressiveness: Testing for conceptual and measurement equivalence across cultures. *Communication Monographs, 60*(4), 256-270.

Reviewer: Ellen A. Hay.

College Self-Expression Scale

Authors/Developers: J. P. Galassi, J. S. DeLo, M. D. Galassi, & S. Bastien. John P. Galassi, Student Counseling Service, West Virginia Univ, Morgantown WV 26506.

Source: Galassi, J. P., DeLo, J. S., Galassi, M. D., & Bastien, S. (1974). The college self-expression scale: A measure of assertiveness. *Behavior Therapy, 5*, 165-171.

Cost: Information not available.

Intended context or behavior: Assertiveness.

Intended academic level: College.

Time required: 15-20 minutes.

Text description: The CSES is a 50-item, self-report measure. It utilizes a 5-point Likert-type scale with 21 positively worded items and 29 negatively worded items. The scale attempts to measure three aspects of assertiveness.

Scoring description: A total score for the scale is obtained by summing all positively worded items, and reverse scoring and summing all negatively worded items. Low scores are indicative of a generalized nonassertive response pattern. Mean scores ranging from 117.91 to 133.00 were reported for various populations studied in this article.

Norm/criterion reference populations: College students and student teachers (n = 200) were used in norm referencing this measure.

Culture and gender bias: Males achieved slightly higher scores than females.

Validity

Predictive: Low to moderate correlations were reported between CSES scores and rating of assertiveness done by the supervisors of the student teachers.

Face/content: The article discusses this conceptualization of assertiveness.

Construct: The CSES correlated with several appropriate subscales on the 300-item Adjective Check List.

Reliability

Test/retest: Coefficients of .89-.90 were established for various groups taking the measure on a test-retest basis.

Evaluative Reactions

Practicality and usefulness: CSES would be helpful in measuring changes in student perceptions of assertiveness.

Clarity of instructions: Directions for completing the measure are very clear.

Standardization in administration and evaluation: Procedures are highly standardized.

Range of responses or degrees of competence: Means are included in the article. Interpretative ranges are not given.

Primary strengths: This is a quick, efficient means of assessing assertiveness.

Primary weaknesses: There seemed to be some question about one's self-perception of assertiveness in relation to the perceptions of others.

Overall adequacy: It may be helpful in showing development as a result of instruction in this area.

Additional References

Galassi, J. P., & Periot, A. R. (1992). What you should know about behavioral assessment—Comment/Reply. *Journal of Counseling and Development, 70*(5), 624-631.

Kirschner, S. M., & Galassi, J. P. (1983). Person, situational, and interactional influences on assertive behavior. *Journal of Counseling Psychology, 30*(3), 355-360.

Reviewer: Ellen A. Hay.

Conversational Sensitivity Scale

Authors/Developers: J. A. Daly, A. L. Vangelisti, & S. M. Daughton. John A. Daly, Dept of Speech Communication, Univ of Texas, Austin TX 78712.

Source: Daly, J. A., Vangelisti, A. L., & Daughton, S. M. (1987). The nature and correlates of conversational sensitivity. *Human Communication Research, 14*, 167-202.

Cost: Information not available.

Intended context or behavior: Interaction/Interpersonal abilities.

Intended academic level: College.

Time required: 15-20 minutes.

Text description: This 36-item self-report is designed to measure the individual's competence in conversational situations. Conversational sensitivity includes eight major components such as perceptiveness, memory, tact, and listening.

Scoring description: Scoring information will need to be obtained from the senior author. In one study conducted using this measure, subjects with a mean score of 109.21 were classified as the low sensitivity group and a mean score of 139.33 as the high sensitivity group.

Norm/criterion reference populations: Various groups of undergraduate students were used in studies of this instrument.

Culture and gender bias: No differences were found between the scores of males and females.

Validity

Predictive: Scores on the CSM did relate to certain behavioral correlates.

Concurrent: Positive correlations were reported between the CSM and Riggio's Social Skills Inventory.

Face/content: The theoretical underpinnings of the conversational sensitivity construct were explained in the article.

Construct: Measures of various communication related variables such as apprehension, empathy, loneliness, self-monitoring, assertiveness, self-esteem, and rhetorical sensitivity were appropriately related to conversational sensitivity.

Reliability

Internal consistency: Alpha coefficients above .80 were reported.

Evaluative Reactions

Practicality and usefulness: With more elaboration, this measure might be used on a pre/post basis to measure development of interpersonal abilities.

Range of responses or degrees of competence: High and low sensitivity groups can be identified.

Primary strengths: The underlying factors identified in this measure relate to many of the concepts and abilities that we consider in interpersonal instruction. This might be one way to measure such instruction.

Primary weaknesses: The measure is designed more to define and explain concept rather than to measure the development of ability.

Overall adequacy: This offers a viable alternative as an assessment instrument and model.

Additional References

Daly, J. A. (1992). Measures of communication-related personality variables. *Communication Education, 41*(4), 405-414.

Stamp, G. H., Vangelisti, A. L., & Daly, J. A. (1992). The creation of defensiveness in social interaction. *Communication Quarterly, 40*(2), 177-190.

Vangelisti, A. L., & Daly, J. A. (1989). Correlates of speaking skills in the United States. *Communication Education, 38*(2), 132-143.

Vangelisti, A. L., Knapp, M. L., Daly, J. A. (1990). Conversational narcissism. *Communication Monographs, 57*(4), 251-274.

Reviewer: Ellen A. Hay.

Interaction Involvement Scale

Author/Developer: Donald J. Cegala, Communication Dept Ohio State Univ, Columbus OH 43210.

Source: Cegala, D. J. (1981). Interaction involvement: A cognitive dimension of communicative competence. *Communication Education, 30*, 108-121.

Cost: Free/public domain.

Intended context or behavior: Willingness to become involved in social interactions.

Intended academic level: College.

Time required: 10-15 minutes.

Text description: This self-report asks respondents to consider 18 statements related to their thoughts and behaviors during conversations with others. Subjects respond on a 7-point Likert-type scale as to how each statement pertains to them.

Scoring description: Ratings can be totaled for an overall score. Scores in the upper 20% are considered high in interaction involvement; those in the bottom 20% are considered low. The scale also yields three subscores that indicate the individual's responsiveness, perceptiveness, and attentiveness.

Norm/criterion reference populations: Various groups of undergraduate students have been used in studies of this measure.

Culture and gender bias: Considerable attention has been given to differences between males and females. Specific findings in relation to each of the subscales are discussed in the 1982 article cited below.

Validity

Predictive: Behaviors during interaction can be predicted on the basis of SSI scores, particularly with regard to the perceptiveness factor.

Concurrent: The SSI was correlated with measures of personality types, communication apprehension, and communication competence. Appropriate relationships among these measures were reported.

Face/content: The theoretical background and scale development are discussed in detail.

Construct: Nonverbal correlations of interaction involvement have been reported.

Reliability

Test/retest: A test-retest reliability of .81 was reported.

Internal consistency: Alpha coefficients for each of the three subscales ranged from .75-.83.

Evaluative Reactions

Practicality and usefulness: This measure might be used on a pre/post basis in an interpersonal course to measure the effectiveness of instruction.

Clarity of instructions: Directions for completing the measure are very clear.

Standardization in administration and evaluation: The measure is highly standardized.

Range of responses or degrees of competence: Interpretative ranges are not offered. Distinctions were made based on percentages within the referencing population.

Primary strengths: This measure has undergone extensive validation. It provides a good overall assessment of interpersonal performance that encompasses a wide range of behaviors.

Primary weaknesses: Scoring and interpretation are a bit ambiguous.

Overall adequacy: This measure seems to provide a valid option for assessing a broad spectrum of interpersonal abilities.

Additional References

Cegala, D. J., Savage, G. T., Brunner, C. G., & Conrad, A. B. (1982). An elaboration of the meaning of interaction involvement. *Communication Monographs, 49*, 229-249.

Cegala, D. J., & Waldron, V. R. (1992, Summer). A study of the relationship between communicative performance and conversation participants' thoughts. *Communication Studies, 43*(2), 105-123.

Chen, G. (1990). *Communication adaptability and interaction involvement as predictors of cross-cultural adjustment*. Kingston, RI: University of Rhode Island. (ERIC Document Reproduction Service No. ED 327 907)

Chen, G. (1990, November). *Communication adaptability and interaction involvement as predictors of cross-cultural adjustment*. Paper presented at the meeting of the Speech Communication Association, Atlanta.

Reviewer: Ellen A. Hay.

Interpersonal Communication Competence Scale

Authors/Developers: R. B. Rubin, & M. M. Martin. R. B. Rubin, School of Communication Studies, Kent State Univ, Kent OH 44242; M. M. Martin, Communication Dept, Mid-Michigan Community College, Harrison MI 48625.

Source: Rubin, R. B., & Martin, M. M. (1994). Development of a measure of interpersonal communication competence. *Communication Research Reports, 11*, 33-44.

Cost: Free/public domain.

Intended context or behavior: Interpersonal skills.

Intended academic level: College/Adult.

Time required: 10 minutes.

Text description: On a 5-point Likert-type scale, students respond to 30 statements designed to measure 10 important interpersonal skills. These skills include self-disclosure, empathy, social relaxation, assertiveness, interaction management, altercentrism, expressiveness, supportiveness, immediacy, and environmental control. A 10-item, short form of the measure is also available.

Scoring description: Several items are reversed for scoring. The scale is intended for use in instructional settings to enable students to see the results of interpersonal instruction. Scores should be compared on a pre/post basis.

Norm/criterion reference populations: College students.

Culture and gender bias: Information not available.

Validity

Concurrent: The 10 short-form items were moderately correlated with separate measures of the 10 interpersonal skills. The short form also showed significant positive correlations to measures of cognitive and communication flexibility.

Face/content: Items for the scale were drawn from existing scales or from definitions of the various constructs.

Construct: Extensive review of the literature resulted in the focus upon 10 important interpersonal skills. Factor analysis was used in selecting the pertinent items.

Reliability

Alternate forms: The 30-item and the 10-item versions correlated at .86.

Internal consistency: Alpha coefficients for the 10 three-item subscales ranged from .41-.72. The overall alpha was .86.

Evaluative Reactions

Practicality and usefulness: The measure is based on 10 interpersonal skills that frequently receive attention in interpersonal instruction. As a form of self-assessment, it will illustrate the acquisition of these skills.

Clarity of instructions: The scale is very clear and easy to use.

Standardization in administration and evaluation: The scale should be relatively simple for everyone to follow.

Range of responses or degrees of competence: Interpretative ranges are not offered.

Primary strengths: This is an excellent resource for interpersonal instruction. It will effectively demonstrate to students how their abilities have changed as a result of instruction.

Primary weaknesses: It would be interesting to see if this scale could be adapted to an other-report format so it could be used as an indicator of student achievement.

Overall adequacy: The scale accomplishes its intended objective.

Additional References

Rubin, R. B., Martin, M. M., Bruning, S. S., & Powers, D. E. (1993). Test of a self-efficacy model of interpersonal communication competence. *Communication Quarterly, 41*, 210-220.

Rubin, R. B. (1990). Communication competence. In G. M. Phillips & J. T. Wood (Eds.), *Speech communication: Essays to commemorate the 75th anniversary of the Speech Communication Association* (pp. 94-129). Carbondale: Southern Illinois Press.

Reviewer: Ellen A. Hay.

Interpersonal Communication Satisfaction Inventory

Author/Developer: M. L. Hecht. Dept of Communication, Arizona State Univ, Tempe AZ 85187.

Source: Hecht, M. L. (1978). The conceptualization and measurement of interpersonal communication satisfaction. *Human Communication Review*, *4*, 253-264.

Cost: Free/public domain.

Intended context or behavior: Satisfaction with interpersonal interactions.

Intended academic level: College.

Time required: 15-40 minutes.

Text description: After either participating in an actual 15-minute conversation or recalling a recent conversation, respondents consider 19 statements pertaining to the interaction. Using a 7-point Likert-type scale, the respondents rate each statement as to whether they agree or disagree with it.

Scoring description: A scoring key is included in the study. Ratings for 8 items are reversed and score is totaled for the 19 items. Mean scores and ranges of responses were not reported in this study.

Norm/criterion reference populations: Various groups of 1st-year students enrolled in an introductory communication course at the University of Illinois participated in the different stages of this study.

Culture and gender bias: Information not available.

Validity

Concurrent: High to moderate correlations were reported between this measure and another using nonverbal faces as means of measuring communication satisfaction.

Face/content: Conceptualization and construction of the instrument were reported in the study. Resulting factor analyses were also reported.

Reliability

Alternate forms: : Reliability coefficients for a 16-item version were reported. The 19-item version was judged as superior.

Internal consistency: Coefficients ranging from .90 to .97 were reported for the measure when studied in various contexts.

Evaluative Reactions

Practicality and usefulness: The measure might be appropriate for a pre- and postindicator of changes in interpersonal effectiveness.

Clarity of instructions: Directions are very clear.

Standardization in administration and evaluation: Actual conversation topics were not provided. The topic may affect the behaviors measured.

Primary strengths: The process of identifying and verifying items for this measure was very thorough.

Overall adequacy: The measure may be better suited to research settings than to instructional ones.

Additional References

Hecht, M. L. (1984). An investigation of the effects of sex on self and other on perceptions of communication satisfaction. *Sex Roles*, *10*, 733-741.

Hecht, M. L. (1989). An Afro-American perspective on interethnic communication. *Communication Monographs*, *56*, 385-410.

Martin, J. N., Hecht, M. L., & Larkey, L. K. (1994). Conversational improvement strategies for interethnic communication: African American and European American perspective. *Communication Monographs*, *61*(3), 236-255.

Reviewer: Ellen A. Hay.

Interpersonal Problem Solving Assessment Technique

Authors/Developers: H. Getter, & J. K. Nowinski Dr. Herbert Getter, Professor of Psychology, Univ of Connecticut, U-11A, Storrs CT 06268.

Source: Getter, H. & Nowinski, J. K. (1981). A free response test of interpersonal effectiveness. *Journal of Personality Assessment, 45,* 301-308.

Cost: Information not available.

Intended context or behavior: Interpersonal communication with specific attention to the ability to generate alternatives and select appropriate responses in various problem situations.

Intended academic level: College/Adult.

Time required: 90-120 minutes.

Text description: Respondents are presented with 46 interpersonal problems that are related to authority, social distance, personal requests, and sexual and contractual situations. Free responses to these situations are designed to measure the number of alternatives generated and the response selected.

Scoring description: A scoring manual which is available from the senior author contains detailed instructions for the scoring protocol. The grand total indicates the number of possible scoring alternatives. Responses are also categorized as effective, avoidant, inappropriate, dependent, and unscorable so that a profile can be developed for each respondent.

Norm/criterion reference populations: Undergraduates (n = 72) and young adults undergoing psychotherapy (n = 23) were used for various studies of this instrument.

Validity

Face/content: The theoretical framework is discussed in the article.

Construct: Appropriate correlations were found with assertiveness, aggressiveness and need definition. Clinical subjects were found to be less effective than those who were termed normal.

Reliability

Scoring: High agreement among scorers was reported.

Internal consistency: Reliability coefficients for each of the four scoring categories ranged from .99 to .82.

Evaluative Reactions

Practicality and usefulness: Because the responses are free-form, the measure is more likely to be an accurate reflection of the individual's communication choices.

Overall adequacy: This measure is designed for more diagnostic purposes. It might be a useful model to consider when developing a free-form communication assessment.

Additional References

Information not available.

Reviewer: Ellen A. Hay.

Opener Scale

Authors/Developers: L. C. Miller, J. H. Berg, & R. L. Archer. L. C. Miller, Dept of Psychology, Northern Illinois Univ, DeKalb IL 60115.

Source: Miller, L. C., Berg, J. H., & Archer, R. L. (1983). Openers: Individuals who elicit intimate self-disclosure. *Journal of Personality and Social Psychology, 44*, 1,234-1,244.

Cost: Information not available.

Intended context or behavior: Ability to encourage others to self-disclose.

Intended academic level: College/Adult.

Time required: 5-10 minutes.

Text description: Respondents indicate on a 5-point Likert-type scale the degree to which the 10 items on this self-report measure are applicable to their interactions. The items are designed to tap the "perceived reactions of others, interest in listening to others, and interpersonal skills" (Miller, Berg, & Archer, 1983, p. 1,235).

Scoring description: The ratings are totaled for an overall score. The article reported a mean score for women of 30.68 and a mean score for men of 28.01.

Norm/criterion reference populations: Various groups of undergraduate students at the University of Texas, Austin, were used in developing and testing this measure.

Culture and gender bias: Women were found to score significantly higher than men.

Validity

Predictive: "High openers reported significantly more close friends than did the low openers.... Low disclosers disclose more when their partner is a high opener than when their partner is a low opener.... We demonstrated that high openers also tend to be the target of disclosure in more long-term relationships" (Miller et al., 1983, p. 1,241).

Face/content: The article discussed this instrument in relation to previous self-disclosure theory and research.

Construct: "The Opener Scale seemed related to similar personality measures (such as perspective taking, sociability, shyness), but the magnitude of those relations suggests that the Opener Scale is measuring a distinct trait" (Miller et al., 1983, p. 1,238).

Reliability

Test/retest: With a 6-week interval, a test-retest reliability of .69 was reported.

Internal consistency: A reliability coefficient of .79 was reported.

Evaluative Reactions

Practicality and usefulness: This measure would be useful when discussing self-disclosure. It might also be good as a pre- and postcourse indication of change in interpersonal interactions.

Clarity of instructions: The directions are very clear.

Standardization in administration and evaluation: The measure is highly standardized.

Range of responses or degrees of competence: Mean scores are provided. During research, respondents with scores in the lower one third were designated as low disclosers, and those with scores in the upper third were labeled high disclosers.

Primary strengths: This is a very quick, efficient measure.

Primary weaknesses: This measures a fairly narrow construct, and there may be a temptation to overgeneralize from it.

Overall adequacy: When used in conjunction with other measures, it may be a viable indicator of interpersonal effectiveness.

Additional References

Berg, J. H., & McQuinn, R. D. (1986). Attraction and exchange in continuing and noncontinuing dating relationships. *Journal of Personality & Social Psychology, 50*(5), 942-952.

Collins, N. L., Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin, 116*(3), 457-475.

Miller, L. C., Cooke, L. L. (1992). Should I brag? Nature and impact of positive and boastful disclosure for women and men. *Human Communication Research, 18*(3), 364-399.

Miller, L. C., & Kenny, D. A. (1986). Reciprocity of self-disclosure at the individual and dyadic levels: A social relations analysis. *Journal of Personality & Social Psychology, 50*(4), 713-719.

Miller, L. C., & Roid, G. H. (1988). Factor-analytically derived scales for the Louisville Behavior Checklist. *Journal of Consulting & Clinical Psychology, 56*(2), 302-304.

Reviewer: Ellen A. Hay.

BEST COPY AVAILABLE

Relational Communication Scale

Authors/Developers: J. H. Burgoon, & J. L. Hale. Judee H. Burgoon, Communication Dept, Univ of Arizona, Tucson AZ 85721.

Source: Burgoon, J. H., & Hale, J. L. (1987). Validation and measurement of the fundamental themes of relational communication. *Communication Monographs, 54*, 19-41.

Cost: Free/public domain.

Intended context or behavior: Relational messages in interpersonal interactions.

Intended academic level: College/Adult.

Time required: 30-40 minutes.

Text description: This measure contains 30 statements that relate to the relational themes of immediacy/affect, similarity/ depth, receptivity/trust, composure, formality, dominance, equality, and task orientation. Following an actual or recalled conversation, participants and/or observers rate interaction participants on an 8-point Likert-type scale as to the applicability of each statement in describing the interaction.

Scoring description: Overall scores and interpretive ranges are not provided in this article. Mean scores are included.

Norm/criterion reference populations: Various groups of undergraduates were used in developing and testing this measure.

Culture and gender bias: Male and female participants showed significant differences.

Validity

Predictive: The measure showed appropriate correlations with nonverbal immediate behaviors, eye contact and gaze, vocal qualities, high reward communicators, credibility, and personality.

Face/content: The article includes background information on the development of the construct along with reports of various factor analyses which were performed in developing this instrument.

Reliability

Internal consistency: Reliability for the relational themes ranged from .81 to .42.

Evaluative Reactions

Practicality and usefulness: This measure can be used on a pre/post basis to identify possible changes in interaction patterns.

Clarity of instructions: The article is primarily research focused and does not include definite instructions for using this measure.

Standardization in administration and evaluation: Again because of the research focus, standardized administration and evaluation procedures are not included.

Primary strengths: This measure would provide an especially good beginning for discussing the themes or factors which influence our interactions.

Primary weaknesses: In order to be used in instructional settings, further elaboration is necessary.

Overall adequacy: This measure provides a very comprehensive perspective on interpersonal communication.

Additional References

Burgoon, J. K., & Baesler, E. J. (1991). Choosing between micro and macro nonverbal measurement: Application to selected vocalic and kinesic indices. *Journal of Nonverbal Behavior, 15*(1), 57-78.

Burgoon, J. K., Birk, T., & Pfau, M. (1990). Nonverbal behaviors, persuasion, and credibility. *Human Communication Research, 17*(1), 140-169.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior, 18*(2), 155-184.

Burgoon, J. K., Buller, D. B., Ebesu, A. S., & Rockwell, P. (1994). Interpersonal deception: V. Accuracy in deception detection. *Communication Monographs, 61*(4), 303-325.

Burgoon, J. K., Dillman, L., & Stern, L. A. (1993). Adaptation in dyadic interaction: Defining and operationalizing patterns of reciprocity and compensation. *Communication Theory, 3*(4), 295-316.

Burgoon, J. K., & Kelley, D. L. (1989). The nature of arousal and nonverbal indices. *Human Communication Research, 16*(2), 217-255.

Burgoon, J. K., & Lepoire, B. A. (1993). Effects of communication expectancies, actual communication, and expectancy disconfirmation on evaluations of communicators and their communication behavior. *Human Communication Research, 20*(1), 67-96.

Burgoon, J. K., LePoire, B. A., Beutler, L. E., Bergan, J., & Engle, D. (1992). Nonverbal behaviors as indices of arousal: Extension to the psychotherapy context. *Journal of Nonverbal Behavior, 16*(3), 159-178.

Burgoon, J. K., & Newton, D. A. (1991). Applying a social meaning model to relational message interpretations of conversational involvement: Comparing observer and participant perspective. *Southern Communication Journal, 56*(2), 96-113.

Burgoon, J. K., & Walther, J. B. (1990). Nonverbal expectancies and the evaluative consequences of violations. *Human Communication Research, 17*(2), 232-265.

Burgoon, J. K., Walther, J. B., & Baesler, E. J. (1992). Interpretations, evaluations, and consequences of interpersonal touch. *Human Communication Research, 19*(2), 237-263.

Reviewer: Ellen A. Hay.

Revised Self-Disclosure Scale

Author/Developer: L. R. Wheeless, Dept of Communication Studies, Univ of North Texas, P. O. Box 5266, Denton TX 76203-0266.

Source: Wheeless, L. R. (1978). A follow-up study of the relationships among trust, disclosure, and interpersonal solidarity. *Human Communication Research, 4*, 143-157.

Cost: Free/public domain.

Intended context or behavior: Self-disclosure.

Intended academic level: College/Adult.

Time required: 5 minutes.

Text description: With a target person in mind, students respond on a 7-point Likert-type scale to 31 statements designed to measure five dimensions of self-disclosure. The five dimensions are intent, amount, content, depth, and accuracy. To measure general disclosiveness, the target individual instruction is replaced by more general instructions.

Scoring description: Scores on several items are reversed. No overall interpretative framework for scores is provided.

Norm/criterion reference populations: Information not available.

Culture and gender bias: Information not available.

Validity

Face/content: Items are drawn from a conceptualization of self-disclosure that included eight factors found in the literature.

Predictive: Self-disclosure and trustworthiness are predictors of interpersonal solidarity.

Construct: Various analyses were performed with earlier versions of the scale in developing this five-dimension version. Self-disclosure correlated in an appropriate manner with measures of solidarity, apprehension, and intimacy.

Reliability

Scoring: The scale has produced coefficient alphas ranging from .81 to .91.

Internal consistency: Reliability for the various dimensions ranges from .84 to .91.

Evaluative Reactions

Practicality and usefulness: This is a very easy, quick means of demonstrating the various dimensions of self-disclosure.

Clarity of instructions: Directions are clear and simple.

Standardization in administration and evaluation: One version uses specific targets for self-disclosure. Standardization is more likely with the more general version.

Range of responses or degrees of competence: Interpretation for responses is not provided.

Primary strengths: This instrument has undergone thorough development and would be effective in illustrating to students how the content and amount of disclosure can change in various situations.

Primary weaknesses: It would be helpful to have some general indicators of how scores compare. Concurrent validity could have been established with similar measures.

Overall adequacy: It fulfills its intended objective.

Additional References

Graham, E. E. (1994). Revised self-disclosure scale. In R. B. Rubin, P. Palmgreen, & H. E. Sypher (Eds.), *Communication research measures: A sourcebook* (pp. 322-326). New York: Guilford Press.

Wheeless, L. R. (1976). Self-disclosure and interpersonal solidarity: Measurement, validation, and relationships. *Human Communication Research, 3*, 47-61.

Wheeless, L. R., & Grotz, J. (1976). Conceptualization and measurement of reported self-disclosure. *Human Communication Research, 2*, 338-346.

Wheeless, L. R., & Grotz, J. (1977). The measurement of trust and its relationship to self-disclosure. *Human Communication Research, 3*, 250-257.

Wheeless, L. R., Nesser, K., & McCroskey, J. C. (1986). The relationships of self-disclosure and disclosiveness to high and low communication apprehension. *Communication Research Reports, 3*, 129-134.

Reviewer: Ellen A. Hay.

BEST COPY AVAILABLE

Speech Evaluation Instrument

Authors/Developers: C. J. Zahn, & R. Hopper. C. J. Zahn, Dept of Communication, Bachelor Hall 160, Miami Univ, Oxford OH 45056.

Source: Zahn, C. J., & Hopper, R. (1985). Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology, 4*, 113-123.

Cost: Information not available.

Intended context or behavior: Reactions to various dialects and vocal qualities.

Intended academic level: College/Adult.

Time required: 20-30 minutes.

Text description: After listening to speakers (or reading a transcript) of various dialectics, respondents rate each on 30 qualities (a 22-item version is also available) structured on a 7-point bipolar semantic differential. The items have been factored into three subscales: superiority, attractiveness, and dynamism.

Scoring description: The authors suggest considering each subscale separately and comparing the speakers across subscales. Since scores will depend upon the stimulus, no interpretative ranges are offered.

Norm/criterion reference populations: Various groups of undergraduate students have been used in studies of this measure.

Validity

Predictive: The authors report that various studies have been conducted demonstrating the predictive validity of each subscale.

Concurrent: The authors report that appropriate correlations were found with related measures.

Face/content: During development, factor analyses were used to validate the items in the subscales.

Reliability

Internal consistency: Reliability coefficients for the three subscales ranged from .80 to .90.

Evaluative Reactions

Practicality and usefulness: This instrument might be adapted for intercultural/multicultural courses as an indicator of changes in student sensitivity to diversity.

Clarity of instructions: Directions for completing the instrument are very clear.

Standardization in administration and evaluation: Test administrators develop their own tapes as the stimulus for this measure, so standardization occurs only with a particular context.

Range of responses or degrees of competence: Interpretation of scores is not provided.

Primary strengths: This measure has been successfully used to study reactions to various groups of people. It would be an interesting starting point for discussion of cultural or gender differences.

Primary weaknesses: The instrument seems better suited to research than to classroom use.

Overall adequacy: The semantic differential is very thorough and the three subscales would provide for interesting discussion.

Additional References

Full bibliography is available with the instrument on ComServe.

Gundersen, D. F. (1989). *An application of the SEI short form to public speaking events*. Phoenix: Arizona State University. (ERIC Document Reproduction Service No. ED 306 617)

Gundersen, D. F. (1989, April). *An application of the SEI short form to public speaking events*. Paper presented at the meeting of the Southern States Communication Association, Louisville.

Zahn, C. J. (1989). The bases for differing evaluations of male and female speech: Evidence from ratings of transcribed conversation. *Communication Monographs, 56*(1), 59-74.

Reviewer: Ellen A. Hay.

Verbal Aggressiveness Scale

Authors/Developers: D. A. Infante, & C. J. Wigley III. Dominic Infante, School of Communication Studies, Kent State Univ, Kent OH 42242.

Source: Infante, D. A., & Wigley, C. J. (1986). Verbal aggressiveness: An interpersonal model and measure. *Communication Monographs, 53*, 61-69.

Cost: Free/public domain.

Intended context or behavior: Measures aggressiveness during interpersonal interaction.

Intended academic level: College/Adult.

Time required: 10 minutes.

Text description: Verbal aggressiveness was defined as a personality trait that predisposes an individual to attack the self-concept of another. The 20 statements on this self-report ask students to rate their behaviors as they try to influence others. Ten are worded positively and 10 negatively. Students respond on a 5-point Likert-type scale.

Scoring description: Negative item scores are reversed and the 20 items are totaled. A mean of 49.1 was reported for the referencing population. The mean for males was 51.97. For females it was 46.38.

Norm/criterion reference populations: Undergraduate students (N = 427)

Culture and gender bias: Significant differences were discovered for men and women.

Validity

Predictive: Verbal aggressiveness correlated with the likeliness to engage in aggressive behaviors.

Concurrent: This measure correlated in the predicted direction with several other related measures of hostility, desirability, adequacy, cognitive complexity, and apprehension.

Face/content: Factor analyses and other development procedures are explained.

Construct: Aggressiveness was found to be distinct from argumentativeness.

Reliability

Test/retest: A high (.82) correlation was reported after a 4-week, test/retest study.

Scoring: An alpha coefficient of .81 was reported.

Evaluative Reactions

Practicality and usefulness: This instrument might be effective as a way of measuring the success of study in argumentation and persuasion, particularly when coupled with the Argumentativeness Scale.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Directions for administrating and scoring the instrument are very clear.

Primary strengths: The instrument would be an effective way of introducing aggressiveness in discussions of conflict or argumentation.

Primary weaknesses: Because ranges are not provided, it is difficult to interpret the meaning of particular scores.

Overall adequacy: Although the focus for this instrument is limited, it could be used effectively in conjunction with other assessments of interpersonal communication.

Additional References

Colbert, K. R. (1993). The effects of debate participation on argumentativeness and verbal aggression. *Communication Education, 42*, 206-214.

Colbert, K. R., & Dorff, T. (1991). *The effects of forensics training on verbal aggression and argumentativeness*. Atlanta: Georgia State University. (ERIC Document Reproduction Service No. ED 340 077)

Colbert, K. R., & Dorff, T. (1991, November). *The effects of forensics training on verbal aggression and argumentativeness*. Paper presented at the meeting of the Speech Communication Association, Atlanta.

Infante, D. A., Hartley, K. C., Martin, M. M., Higgins, M. A Bruning, S. D., & Hur, G. (1992a). Initiating and reciprocating verbal aggression: Effects on credibility and credited valid arguments. *Communication Studies, 43*(3), 182-190.

Infante, D. A., Riddle, B. L., Horvath, C. L., & Tumlin, S. A. (1992b). Verbal aggressiveness: Messages and reasons. *Communication Quarterly, 40*(2), 116-126.

Infante, D. A., Myers, S. A., & Buerkel, R. A. (1994). Argument and verbal aggression in constructive and destructive family and organizational disagreements. *Western Journal of Communication, 58*(2), 73-84.

Infante, D. A., & Gorden, W. I. (1991). How employees see the boss: Test of an argumentative and affirming model of supervisors' communicative behavior. *Western Journal of Speech Communication, 55*(3), 294-304.

Infante, D. A., & Rancer, A. S. (1993). Relations between argumentative motivation, and advocacy and refutation on controversial issues. *Communication Quarterly, 41*(4), 415-426.

Remland, M. S., Jones, T. S., & Brown, T. (1994). *Nonverbal communication and conflict: The effects of attribution and predisposition*. West Chester, PA: West Chester University. (ERIC Document Reproduction Service No. ED 372 440)

Remland, M. S., Jones, T. S., & Brown, T. (1994, June). *Nonverbal communication and conflict: The effects of attribution and predisposition*. Paper presented at the meeting of the International Association of Conflict Management, Eugene, OR.

Reviewer: Ellen A. Hay.

BEST COPY AVAILABLE

Class Apprehension about Participation Scale (CAPS)

Author/Developer: M. R. Neer. Dept of Communication Studies, Univ of Missouri, 5100 Rockhill Road, Kansas City MO 64110.

Source: Neer, M. R. (1987). The development of an instrument to measure classroom apprehension. *Communication Education, 36*, 154-166.

Cost: Free/public domain.

Intended context or behavior: Communication apprehension.

Intended academic level: College/Adult.

Time required: 20-30 minutes.

Text description: A self-report inventory of 20 Likert-type statements designed to measure level of classroom apprehension and to identify observable behaviors. The dimension called communication participation consists of 10 items which generally index approach/avoidance behaviors associated with communication apprehension, and the dimension called communication confidence is made up of 10 items which include general nervousness and level of fear of being evaluated during class discussion.

Scoring description: Both instruments have instructions included at the beginning, so they can be administered in large settings if necessary. Items scores are summed for each of the two dimensions, or added together to an overall summary score.

Norm/criterion reference populations: Studies were conducted on 324 undergraduates enrolled in a basic public speaking course. The mean was 52.68, with a standard deviation of 15.61.

Culture and gender bias: Both males and females were included in the validity study.

Validity

Predictive: CAPS scores correlated with self-reports of 11 or 12 instructional practices which help students feel more comfortable about participating in class.

Concurrent: Overall CAPS scores correlated at .71 with the PRCA-24. Each dimension (confidence and participation) correlated at .74 with the PRCA-24.

Reliability

Internal consistency: Alpha reliability for overall score was .94. For confidence = .88, and for the participation dimension = .91.

Evaluative Reactions

Practicality and usefulness: The scales are easily used. It provides information about classroom behaviors and evaluative attitudes which is not available with other instruments.

Clarity of instructions: The directions are clear.

Range of responses or degree of competence: Students' scores have been divided into low, medium, and high apprehensives based on mean deviates.

Primary strengths: It has been compared favorably with the PRCA and is easily administered.

Primary weaknesses: The CAPS scale has been used primarily in research of communication variables rather than as an assessment technique.

Overall adequacy: This has good potential to be developed as an assessment technique.

Additional References

Aitken, J. E., & Neer, M. R. (1992). *The relationship of classroom communication apprehension and motivation to college student questioning*. Kansas City: University of Missouri-Kansas City. (ERIC Document Reproduction Service No. ED 347 598)

Aitken, J. E., & Neer, M. R. (1992, November). *The relationship of classroom communication apprehension and motivation to college student questioning*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Neer, M. R. (1990). *Instructor communication behavior as a factor influencing the class participation of classroom communication apprehensives*. Kansas City: University of Missouri-Kansas City. (ERIC Document Reproduction Service No. ED 324 731)

Neer, M. R. (1990, November). *Instructor communication behavior as a factor influencing the class participation of classroom communication apprehensives*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Reviewer: Judith M. Dallinger.

Communication Anxiety Inventory

Authors/Developers: S. Booth-Butterfield, & M. Gould. S. Booth-Butterfield, Communication Studies Dept, West Virginia Univ, Morgantown WV 26506.

Source: Booth-Butterfield, S., & Gould, M. (1986). The communication anxiety inventory: Validation of state and context communication apprehension. *Communication Quarterly, 34*, 194-205.

Cost: Free/public domain.

Intended context or behavior: Communication apprehension.

Intended academic level: College/Adult.

Time required: 20-30 minutes.

Text description: Two instruments are available: one called Form Trait and one called Form State. Form Trait includes 21 4-point, Likert-type items assessing communication apprehension in three specific contexts: dyadic, small group, and public speaking. Form State is made up of 20 4-point, Likert-type items.

Scoring description: Both instruments have instructions included at the beginning, so they can be administered in large settings if necessary. Scores for state apprehension, trait apprehension, and trait apprehension in particular contexts are calculated by summing appropriate items.

Norm/criterion reference populations: A series of studies on the Form Trait measure with a total of 754 high school and college students resulted in a mean of 46.90 with a standard deviation of 10.495. That sample resulted in means and standard deviations for trait apprehension in particular contexts of (a) interpersonal: mean = 12.996, sd = 3.204; small group: mean = 15.033, sd = 4.235; and (b) public speaking: mean = 18.874, sd = 5.174. On a sample of 163 undergraduate students, the mean for Form State was 44.06 with a standard deviation of 11.07.

Culture and gender bias: Both males and females were included in the scale development sample.

Validity

Predictive: Students were involved in an experimental condition in which they participated in instances of communication for each of the three contexts. Following their participation, they completed the CAI scales and other state anxiety measures. Results showed that scores on the different sections of Form Trait correlated best with state apprehension scores based on experience in the respective contexts.

Reliability

Internal consistency: Alpha reliability for Form Trait: .898; Interpersonal Trait: .654; Small Group Trait: .846; Public Speaking Trait: .887; Form State: .912.

Evaluative Reactions

Practicality and usefulness: The scales are easily used.

Clarity of instructions: Directions are clear.

Range of responses or degree of competence: Not enough data have been collected to stabilize norms.

Primary strengths: It is easily administered and conceptually sound.

Primary weaknesses: The CAI has been used primarily in research of communication variables rather than as an assessment technique.

Overall adequacy: This has good potential to be developed as an assessment technique.

Additional References

Booth-Butterfield, S. (1988). Inhibition and student recall of instructional messages. *Communication Education, 37*(4), 312-324.

Booth-Butterfield, S. (1991). Ethical issues in the treatment of communication apprehension and avoidance. *Communication Education, 40*, 172-179.

Booth-Butterfield, M., & Booth-Butterfield, S. (1993). The role of cognitive "performance orientation" in communication anxiety. *Communication Quarterly, 41*, 198-209.

Rekart, D. M., Begnal, C. F., & Mino, M. (1990). *Speech modification training for reticent speakers: A comparison of three methods*. University Park, PA: The Pennsylvania State University. (ERIC Document Reproduction Service No. ED 321 314)

Rekart, D. M., Begnal, C. F., & Mino, M. (1990, April). *Speech modification training for reticent speakers: A comparison of three methods*. Paper presented at the meeting of the Eastern Communication Association, Philadelphia.

Reviewer: Judith M. Dallinger.

BEST COPY AVAILABLE

Personal Report of Communication Apprehension 24 (PRCA-24)

Author/Developer: J. C. McCroskey. Communication Studies, West Virginia Univ, Morgantown WV 26506-6293.

Source: McCroskey, J. C. (1986). *An Introduction to Rhetorical Communication* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Cost: Free/public domain.

Intended context or behavior: Communication apprehension.

Intended academic level: College/Adult.

Time required: 15-25 minutes.

Text description: Likert-type items index communication apprehension in each of four different contexts: public speaking, speaking in small groups, speaking in meetings, and speaking in dyads. Six items, including three positively worded and three negatively worded, address each context. Context scores can be used separately or summed for trait communication apprehension.

Scoring description: The instrument has instructions included at the beginning, so it can be administered in large settings if necessary. Scores for each context and a score for overall CA are developed by combining scores on the various items.

Norm/criterion reference populations: Mean score is 65.60, with a standard deviation of 15.30, based on 25,000 students from 52 colleges and universities.

Culture and gender bias: The factor structure did not hold for a sample of Puerto Rican students, so caution should be used when using this questionnaire with intercultural students.

Validity

Predictive: The PRCA-24 accounts for 50% of the variance in the RAS assertiveness scale.

Face/content: Studies show that 94-96% of the total PRCA-24 score is accounted for by any three of the four contexts, which indicates that the scale is measuring generalized trait-like communication apprehension.

Reliability

Internal consistency: Alpha reliability normally is .93 to .95.

Evaluative Reactions

Practicality and usefulness: The scale is easy to administer and analyze.

Clarity of instructions: Directions are clear.

Standardization in administration and evaluation: Norms are stable and have been developed on a large sample.

Primary strengths: A great deal of data have been collected to establish norms.

Primary weaknesses: The scale has been used often, but has not been used as much as an assessment instrument.

Overall adequacy: This is useful as an assessment instrument.

Additional References

Beatty, M., & Andriate, G. S. (1985). Communication apprehension and general anxiety in the prediction of public speaking anxiety. *Communication Quarterly, 33*, 174-184.

Bourhis, J. S., & Stubbs, J. (1991). *Communication apprehension and learning styles*. Springfield, MO: Southwest Missouri State University. (ERIC Document Reproduction Service No. ED 335 713)

Bourhis, J. S., & Stubbs, J. (1991, April). *Communication apprehension and learning styles*. Paper presented at the meeting of the Central States Communication Association, Chicago.

DeWine, S., & Pearson, J. C. (1985). *The most frequently used self-report instruments in communication*. Athens: Ohio University. (ERIC Document Reproduction Service No. ED 260 479).

DeWine, S., & Pearson, J. C. (1985, May). *The most frequently used self-report instruments in communication*. Paper presented at the meeting of the International Communication Association, Honolulu.

Fredricks, R., Stinson, C., & Soukup, P. A. (1993). *Communication apprehension among adult children of alcoholics*. Santa Clara, CA: Santa Clara University. (ERIC Reproduction Service No. ED 364 923)

Fredricks, R., Stinson, C., & Soukup, P. A. (1993, November). *Communication apprehension among adult children of alcoholics*. Paper presented at the meeting of the Speech Communication Association, Miami Beach.

Kelly, L, & Keaten, J. (1991). *Does the Penn State Reticence Program work?: A comparison of special and regular options of a basic speech communication course*. Erie, PA: Penn State University. (ERIC Document Reproduction Service No. ED 344 269)

Kelly, L, & Keaten, J. (1991, November). *Does the Penn State Reticence Program work?: A comparison of special and regular options of a basic speech communication course*. Paper presented at the meeting of the Speech Communication Association, Atlanta.

Levine, T. R., & McCroskey, J. C. (1990). Measuring trait communication apprehension: A test of rival measurement models of the PRCA-24. *Communication Monographs, 57*, 62-71.

McCroskey, J. C. (1983). *Cross-situational consistency of the PRCA: Another view*. Morgantown, WV: West Virginia University. (ERIC Document Reproduction Service No. 226 383)

McCroskey, J. C. (1983, February). *Cross-situational consistency of the PRCA: Another view*. Paper presented at the meeting of the Western Speech Communication Association, Albuquerque.

McCroskey, J. C, Beatty, M. J., Kearney, P., & Plax, T. G. (1985). The content validity of the PRCA-24 as a measure of communication apprehension across communication contexts. *Communication Quarterly, 33*, 165-173.

Pearson, J. C., & Turner, L. H. (1984). *The personal report of communication apprehension: Predictive validity and behavioral correlates*. Athens: Ohio University. (ERIC Document Reproduction Service No. 243 157)

Pearson, J. C., & Turner, L. H. (1984, March). *The personal report of communication apprehension: Predictive validity and behavioral correlates*. Paper presented at the meeting of the Eastern Communication Association, Philadelphia.

Reviewer: Judith M. Dallinger.

Receiver Apprehension Test

Author/Developer: L. R. Wheeless. Dept of Communication Studies, Marshall Univ, 400 Hall Greer Blvd., Huntington WV 25755-2632.

Source: Wheeless, L. R. (1975). An investigation of receiver apprehension and social context dimensions of communication apprehension. *The Speech Teacher, 24*, 261-268.

Cost: Free/public domain.

Intended context or behavior: Communication apprehension.

Intended academic level: College/Adult.

Time required: 15-25 minutes.

Text description: Twenty Likert-type items index receiver apprehension on 5-point scale.

Scoring description: The instrument has instructions included at the beginning, so it can be administered in large settings if necessary. A score is developed by combining scores on all the items.

Norm/criterion reference populations: The original analysis resulted in a mean of 46.93, with a standard deviation of 12.67, based on a sample of 324 college students.

Culture and gender bias: Both males and females were included in the scale development sample.

Validity

Predictive: The RAT scores have correlated at low levels with STAI-State scores in reaction to specific reading exercises. The low correlation is appropriate because it indicates that RAT scores are more trait than state.

Face/content: Items were developed to have face validity.

Construct: RAT scores are relatively independent of communication apprehension as measured by the PRCA. The Pearson correlation for the two scales was .20.

Reliability

Internal consistency: Alpha reliability was found to be .80. Split-half reliability was .91.

Evaluative Reactions

Practicality and usefulness: The scale is easily used.

Clarity of instructions: The directions are clear.

Range of responses or degree of competence: Not enough data has been collected to stabilize norms.

Primary strengths: It is easily administered and conceptually sound.

Primary weaknesses: The RAT has been used primarily in research of communication variables rather than as an assessment technique.

Overall adequacy: This has good potential to be developed as an assessment technique.

Additional References

Beatty, M. J., Behnke, R. R., & Henderson, L. S. (1980). An empirical validation of the receiver apprehension test as a measure of trait listening anxiety. *The Western Journal of Speech Communication, 44*, 132-136.

Bock, D. G., & Bock, E. H. (1984). The effects of positional stress and receiver apprehension on leniency errors in speech evaluation: A test of the rating error paradigm. *Communication Education, 33*, 337-341.

Preiss, R. W., & Karssen, J. D. (1988). *Receiver apprehension in educational settings: A typology of outcomes and comprehensive bibliography*. Tacoma, WA: University of Puget Sound. (ERIC Document Reproduction Service No. 298 581)

Preiss, R. W., & Rindo, J. (1985). *The information processing consequences of receiver apprehension*. Tacoma, WA: University of Puget Sound. (ERIC Document Reproduction Service No. 263 186)

Roberts, C. V. (1985). *Preliminary research employing the Watson-Barker Listening Test: A validation of the instrument*. Johnson City, TN: East Tennessee State University. (ERIC Document Reproduction Service No. 257 145)

Roberts, C. V. (1985, March). *Preliminary research employing the Watson-Barker Listening Test: A validation of the instrument*. Paper presented at the meeting of the International Listening Association, Orlando.

Roberts, C. V. (1986, April). *The question of validity and application: How do we know how well we are listening*. Paper presented at the meeting of the Eastern Communication Association, Atlantic City.

Reviewer: Judith M. Dallinger.

Unwillingness-to-Communicate Scale

Author/Developer: J. K. Burgoon. Communication Dept, Speech #25, Univ of Arizona, Tucson AZ 85721.

Source: Burgoon, J. K. (1976). The unwillingness-to-communicate scale: Development and validation. *Communication Monographs, 43,* 60-69.

Cost: Free/public domain.

Intended context or behavior: Communication apprehension/avoidance.

Intended academic level: College/Adult.

Time required: 20-30 minutes.

Text description: "The UCS, a Likert-type scale, has 20 items measuring two dimensions of communication reticence: approach-avoidance (AA) and reward (R). The AA dimension represents an individual's tendency to avoid or participate in interpersonal and small group interactions. In tapping into one's behavioral and anxiety predispositions, it shows a kinship to the PRCA. The R dimension, by contrast, reflects attitudes toward communication—whether one considers it a valuable, honest, and personally rewarding enterprise or feels socially isolated and regards communication as a deceptive, manipulative, or unprofitable activity" (Burgoon & Koper, 1984, p. 608).

Scoring description: Information not available.

Norm/criterion reference populations: Various groups of students at West Virginia University were used to validate this instrument.

Culture and gender bias: Information not available.

Validity

Face/content: Conceptualization of the construct and resulting factor analyses are reported in the initial article. Subsequent research has further validated the dimensions of the scale.

Construct: Moderate positive correlations were reported with measures of communication apprehension, anomia, and alienation. These variables served as a basis for the original design of the instrument. Peer ratings and observer ratings following a group decision-making activity also were significantly correlated with the two factors in the scale.

Reliability

Internal consistency: "Coefficient alpha and split-half reliabilities have ranged from .81 to .88 for the approach-avoidance factor and .60 to .82 for the reward factor" (Burgoon & Hale, 1983, p. 306).

Evaluative Reactions

Practicality and usefulness: This scale conceptualizes reticence in a more complex manner, and therefore may be helpful in identifying students who are significantly apprehensive. It can be used as a pre- and post-course measure of apprehension in a variety of contexts.

Primary strengths: Research using this scale has extended consideration of reticence to the verbal and nonverbal dimensions. Conceptualizing reticence as both an approach-avoidance factor and a reward factor provides a greater understanding of what might be contributing to a student's apprehension.

Primary weaknesses: Administration procedures and scoring information are not included in published accounts of the scale and thus could not be evaluated.

Additional References

Burgoon, J. K., & Hale, J. L. (1983a). A research note on the dimensions of communication reticence. *Communication Quarterly, 31,* 238-248.

Burgoon, J. K., & Hale, J. L. (1983b). Dimensions of communication reticence and their impact on verbal encoding. *Communication Quarterly, 31,* 302-312.

Burgoon, J. K., & Koper, R. J. (1984). Nonverbal and relational communication associated with reticence. *Human Communication Research, 10,* 601-616.

Reviewer: Ellen A. Hay.

Hall Conflict Management Survey

Author/Developer: J. Hall.

Address: Information not available.

Source: Hall, J. (1986). *Conflict management survey: A survey of one's characteristic reaction to and handling of conflict between himself and others.* Conroe, TX: Teleometrics.

Cost: Scale is copyrighted; costs are available from the author.

Intended context or behavior: Conflict Style.

Intended academic level: College/Adult.

Time required: 30-45 minutes.

Text description: The scale is a self-assessment of three statements each for the personal, interpersonal, small group, and intergroup contexts. The 12 statements when taken together represent the overall profile. Each of the 12 statements has five possible reactions that represent the five conflict style preferences: win-lose, yield-lose, lose-leave, compromise, and synergistic.

Scoring description: Five raw scores are calculated for each conflict style from the rating numbers assigned to the 12 statements representing each style.

Norm/criterion reference populations: A *t*-score mean for the overall scale is 50, with a standard deviation of 10. *t*-score means for each context are 10, with standard deviations of 2.

Culture and gender bias: Studies showed no differences between males and females.

Validity

Predictive: Scores correlated fairly strongly with social desirability scores.

Concurrent: As reported by Hall, scores correlate with various personality measures.

Construct: Scores correlate with the Thomas-Kilman MODE instrument on all five modes.

Reliability

Internal consistency: synergistic = .73, compromise = .45, yield-lose = .57, win-lose = .61, and lose-leave = .39.

Evaluative Reactions

Clarity of instructions: It has been reported to be more difficult to administer than other conflict instruments.

Primary strengths: The instrument has been used a great deal in research settings. It has been widely used for training in business and industry.

Additional Reference:

Shockley-Zalabak, P. (1988). Assessing the Hall Conflict Management Survey. *Management Communication Quarterly, 1,* 302-320.

Reviewer: Judith M. Dallinger.

Putnam-Wilson Organizational Communication Conflict Instrument (OCCI)

Author/Developer: L. Putnam. Dept of Speech Communication, Texas A & M Univ, College Station TX 77843-4111.

Source: Putnam, L. L., & Wilson, C. (1982). Communicative strategies in organizational conflict: Reliability and validity of a measurement scale. *Communication Yearbook, 6*, 629-652.

Cost: Information not available.

Intended context or behavior: Conflict Style.

Intended academic level: College/Adult.

Time required: 10-15 minutes.

Text description: The scale includes 30 items that assess conflict strategies of respondents in hypothetical situations. Items are 7-point Likert-type indicating frequencies of behaviors for three dimensions: nonconfrontation, solution orientation, and control.

Scoring description: Scores for dimensions are summed.

Norm/criterion reference populations: Frequency means for 360 participants from six studies range from 2.96 to 5.05 for solution orientation, 1.95 to 3.94 for control, and 1.91 to 3.76 for nonconfrontation.

Culture and gender bias: Some sex differences have been found. Differences occur for high and low context culture. Differences occur between white and black samples.

Validity

Concurrent: Scores correlate with communicator style.

Face/content: Verbal and nonverbal behavior strategies are included.

Reliability

Internal consistency: Solution orientation = .79 to .88, control = .71 to .84, and nonconfrontation = .83 to .93.

Evaluative Reactions

Practicality and usefulness: It is easily used.

Clarity of instructions: Instructions are clear.

Primary strengths: Focuses on communication behavior and includes both verbal and nonverbal. Useful for organizational settings.

Primary weaknesses: Test/retest reliability is not sufficient to suggest the instrument's use as a training measure.

Additional References

McDowell, E. E. (1990). *A study of the relationship between Willingness to Communicate and Preferred Conflict Strategy: Implications for teaching communication and conflict*. St. Paul: University of Minnesota. (ERIC Document Reproduction Service No. ED 322 561)

McDowell, E. E. (1990, June). *A study of the relationship between Willingness to Communicate and Preferred Conflict Strategy: Implications for teaching communication and conflict*. Paper presented at the meeting of the International Communication Association, Dublin, Ireland.

Putnam, L. L., & Folger, J. P. (1988). Communication, conflict and dispute resolution: The study of interaction and the development of conflict theory: An introduction to the special issue. *Communication Research, 15*, 49-59.

Ting-Toomey, S. (1984). *Conflict management styles in black and white subjective culture*. New Brunswick, NJ: Rutgers University. (ERIC Document Reproduction Service No. ED 243 166)

Ting-Toomey, S. (1984, May). *Conflict management styles in black and white subjective culture*. Paper presented at the annual meeting of the International Communication Association, San Francisco.

Wilson, S. R., & Waltman, M. S. (1988). Assessing the Putnam-Wilson Organizational Conflict Instrument (OCCI). *Management Communication Quarterly, 1*, 367-388.

Reviewer: Judith M. Dallinger.

Rahim Organizational Conflict Inventory-II (ROCI-II)

Author/Developer: M. A. Rahim. 3803 E. Bayshore Rd, Palo Alto CA 94303. 415-969-8901.

Source: Rahim, M. A. (1983). *Rahim organization conflict inventories: Professional manual.* Palo Alto, CA: Consulting Psychologists Press.

Cost: Inventory Form A, B, or C (25 units): $25.20; Answer Sheets (25 units): $25.00; and Manual: $22.05.

Intended context or behavior: Conflict style.

Intended academic level: College/Adult.

Time required: 8 minutes.

Text description: The scale includes twenty-eight, 5-point Likert-type items which measure five styles of conflict management. The five styles include integrating, obliging, dominating, avoiding, and compromising. There are three forms: Form A for boss, Form B for subordinate, and Form C for peer.

Scoring description: Scores for dimensions are averaged.

Norm/criterion reference populations: Data from 1,219 managers result in means of 4.18 for integrating, 3.60 for obliging, 3.27 for dominating, 2.89 for avoiding, and 3.51 for compromising on Form A.

Culture and gender bias: Information not available.

Validity

Predictive: Dominating scores negatively predicted concession in a negotiation task.

Concurrent: There are small significant correlations between the five styles and six behavioral and cognitive reactions to inequality.

Face/content: Face validity is evident for the five styles.

Reliability

Test/retest: Integrating = .83, obliging = .81, dominating = .76, avoiding = .79, and compromising = .60.

Internal consistency: Integrating = .65-.89, obliging = .68-.87, dominating = .66-.81, avoiding = .61-.86, and compromising = .50-74.

Evaluative Reactions

Practicality and usefulness: It has been used by trainers in organizations.

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Means are available.

Primary strengths: It has a good theoretical basis. Good psychometric properties.

Additional References

Rahim, M. A. (1983). Measurement of organizational conflict. *The Journal of General Psychology, 109*, 189-199.

Rahim, M. A., & Magner, N. R. (1994). Convergent and discriminant validity of the Rahim organizational conflict inventory II. *Psychological Reports, 74*, 35-38.

Short, P. M., & Johnson, P. E. (1994). *Exploring the links among teacher empowerment, leader power, and conflict.* University Park: The Pennsylvania State University. (ERIC Document Reproduction Service No. 372 044)

Van Epps, P. D., & Thompson, B. (1990). *Conflict management styles predicted by managerial frames of reference.* New Orleans: University of New Orleans. (ERIC Document Reproduction Service No. 320 952)

Van Epps, P. D., & Thompson, B. (1990, June). *Conflict management styles predicted by managerial frames of reference.* Paper presented at the International Conference of the International Association for Conflict Management, Vancouver, CA.

Weider-Hatfield, D. (1988). Assessing the Rahim Organizational Conflict Inventory-II (ROCI-II). *Management Communication Quarterly, 1*, 350-366.

Reviewer: Judith M. Dallinger.

Ross-DeWine Conflict Management Message Style (CMMS)

Authors/Developers: R. Ross, & S. DeWine. Roseanna Ross, St. Cloud State Univ, St. Cloud MN 56301.

Source: Ross, R. R., & DeWine, S. (1988). Assessing the Ross-DeWine Conflict Management Message Style (CMMS). *Management Communication Quarterly, 1*, 389-413.

Cost: Information not available.

Intended context or behavior: Conflict style.

Intended academic level: College/Adult.

Time required: 10-15 minutes.

Text description: A self-report scale instrument used to measure conflict styles. Eighteen items include six each for three dimensions: self, issues, and other.

Scoring description: Scores for dimensions are summed.

Norm/criterion reference populations: Scores range from 6-30 for each dimension.

Culture and gender bias: No biological sex differences were found. Some psychological sex differences occur. Students from the U.S. and Denmark reported different scores.

Validity

Predictive: Self-reports correlated with reports by friends.

Concurrent: Scores correlate between conflict styles and interpersonal needs at moderate levels.

Face/content: Trainers report that items reflect their choice of styles.

Construct: Correlated with Putnam-Wilson OCCI, with Hall Conflict Scale, and with Bell and Blakeney.

Reliability

Test/retest: Self = .69, issue = .65, other = .63.

Internal consistency: Alpha reliability: .76 for self, .78 for issues, and .73 for other.

Evaluative Reactions

Practicality and usefulness: It is easily used and relatively short.

Clarity of instructions: Instructions are clear.

Primary strengths: Focuses on communication behaviors and has been used in training.

Primary weaknesses: Has not been used as an assessment measure.

Overall adequacy: It has potential to be a good assessment instrument.

Additional References

Ross, R., & DeWine, S. (1984). *Interpersonal needs and communication in conflict: Do soft words win hard hearts?* St. Cloud, MN: St. Cloud State University. (ERIC Document Reproduction Service No. 250 750)

Ross, R., & DeWine, S. (1984, November). *Interpersonal needs and communication in conflict: Do soft words win hard hearts?* Paper presented at the meeting of the Speech Communication Association, Chicago.

Reviewer: Judith M. Dallinger.

Thomas-Kilman Conflict MODE Survey

Authors/Developers: K. W. Thomas, & R. H. Kilman 914-351-4735.

Source: Thomas, K. W., & Kilman, R. H. (1974). *Thomas-Kilman conflict MODE instrument*. Tuxedo, NY: XICOM.

Cost: $5.50 each. Quantity discounts: 100-249 = $5.15 each; 250-499 = $5 each; over 500 = $4.50 each.

Intended context or behavior: Conflict Style.

Intended academic level: College/Adult.

Time required: 10-20 minutes.

Text description: The scale includes 30 pairs of items, which result in 60 statements. For each pair, the respondent chooses which of the items best describes the way s/he acts in conflict situations. The scale measures behavioral intentions.

Scoring description: The score for a given mode is the number of times the mode is chosen in the scale. The range is 0-12 for each of five modes. The interpretation section presents uses and recommendations for each mode.

Norm/criterion reference populations: Data were collected from 399 middle and upper managers.

Validity

Predictive: The MODE predicted actual behaviors for avoiders. Scores correlated fairly strongly with social desirability scores.

Concurrent: Collaboration and compromise are correlated with communication satisfaction and have been found to be independent of one another.

Face/content: Face, intrinsic, representativeness: Items representing cooperativeness and assertiveness dimensions are included.

Construct: Scores correlate with the CCI conflict instrument as expected.

Reliability

Internal consistency: Modes: avoid = .62, compete = .71, compromise = .58, accommodate = .43, and collaborate = .65.

Evaluative Reactions

Clarity of instructions: Instructions are clear.

Standardization in administration and evaluation: Quartiles have been defined for each mode. Administration is easy.

Primary strengths: The instrument has been used in a great deal of research.

Primary weaknesses: It correlated fairly highly with social desirability. Norms are based on a small sample.

Additional References

Berryman-Fink, C., & Brunner, C. (1985). *The effects of sex of actor and target on interpersonal conflict management styles*. Cincinnati: University of Cincinnati. (ERIC Document Reproduction Service No. 264 629)

Berryman-Fink, C., & Brunner, C. (1985, November). *The effects of sex of actor and target on interpersonal conflict management styles*. Paper presented at the meeting of the Speech Communication Association, Denver.

Womack, D. F. (1988). Assessing the Thomas-Kilman Conflict MODE Survey. *Management Communication Quarterly, 1*, 321-349.

Reviewer: Judith M. Dallinger.

Basic Course Communication Competency Measure

Authors/Developers: W. S. Z. Ford, & A. D. Wolvin. Dept of Speech Communication, Univ of Maryland, 2110 Skinner, College Park MD 20742.

Source: Ford, W. S. Z., & Wolvin, A. D. (1993). The differential impact of a basic communication course on perceived communication competencies in class, work, and social contexts. *Communication Education, 42*, 215-223.

Cost: Information not available.

Intended context or behavior: Abilities and knowledge commonly considered in the hybrid basic communication course.

Intended academic level: College.

Time required: 20-25 minutes.

Text description: This self-report measure is comprised of 24 statements relating to communication abilities (public speaking, interpersonal communication, interviewing, group communication, listening, and self-confidence) frequently considered in the beginning college communication course. On a 7-point Likert-type scale, respondents indicate how well each ability was performed in three contexts: class, social/family, and work.

Scoring description: The ratings are totaled for an overall score. The article provides mean scores for each of the three contexts and each of the 24 statements.

Norm/criterion reference populations: Students (N = 344) enrolled in a basic communication course at a large university served as the referencing population.

Validity

Construct: Students who completed the measure on a pre- and postcourse basis demonstrated significantly higher scores after the completion of the course. These results would be more meaningful were a posttest, control-group design also used.

Reliability

Internal consistency: Cronbach alpha coefficients ranged from .93 to .95 on the three context subscales.

Evaluative Reactions

Practicality and usefulness: Used as it was in this study, this measure provides a good indicator of perceived change in communication abilities.

Clarity of instructions: Directions are clear. The article includes information on structuring the measure to avoid response fatigue.

Standardization in administration and evaluation: The measure is highly standardized.

Range of responses or degrees of competence: Mean scores were reported, but other interpretative data were not provided in this article.

Primary strengths: This measure focuses on the most highly enrolled communication course and may be one means of demonstrating the effectiveness of such a course.

Primary weaknesses: Since the measure is a self-report, the outcomes may not correspond with the actual development of these communication abilities.

Overall adequacy: It is an efficient, cost-effective means of gathering information from students on their communication perceptions.

Additional Reference:

Ford, W. S. Z., & Wolvin, A. D. (1992). Evaluation of a basic course in speech communication. In L. W. Hugenberg (Ed.), *Basic communication course annual* (Vol. 4, pp. 35-47). Boston: American Press.

Reviewer: Ellen A. Hay.

Communicative Adaptability Scale

Authors/Developers: R. I. Duran, & L. Wheeless. R. Duran, Communication Dept, Univ of Hartford, Hartford CT 06117.

Source: Duran, R. I. (1983). Communicative adaptability: A measure of social communicative competence. *Communication Quarterly, 31*, 320-326.

Cost: Free/public domain.

Intended context or behavior: Communicator Style.

Intended academic level: College/Adult.

Time required: 20-30 minutes.

Text description: A 30-item, 5-point, Likert-type scale that includes six dimensions: social composure, wit, appropriate disclosure, articulation, social expertise, and social confirmation.

Validity

Predictive: Groups of subscales correlate as expected with communication apprehension and with self-esteem.

Reliability

Internal consistency: For the scale as a whole = .81. Factors reliabilities range from .70 to .89.

Evaluative Reactions

Practicality and usefulness: The scale is easy to administer.

Clarity of instructions: The directions are clear.

Standardization in administration and evaluation: Scoring instructions are clear enough to lead to standardization.

Range of responses or degree of competence: Large sample norms are not available.

Primary strengths: It focuses on an important aspect of communication.

Primary weaknesses: It was not been developed specifically as an assessment tool.

Overall adequacy: This has potential to be developed as an assessment technique.

Additional References

Chen, G. M. (1990). *Communication adaptability and interaction involvement as predictors of cross-cultural adjustment*. Kingston, RI: University of Rhode Island. (ERIC Document Reproduction Service No. 327 907)

Chen, G. M. (1990, November). *Communication adaptability and interaction involvement as predictors of cross-cultural adjustment*. Paper presented at the meeting of the Speech Communication Association, Chicago.

Downs, C. W., Archer, J., McGrath, J., & Stafford, J. (1988). An analysis of communication style instrumentation. *Management Communication Quarterly, 1*, 543-571.

Duran, R. L. (1982). *Communicative competence a question of context: It depends....* Hartford, CT: University of Hartford. (ERIC Document Reproduction Service No. 219 809)

Duran, R. L. (1982, May). *Communicative competence a question of context: It depends....* Paper presented at the meeting of the Eastern Communication Association, Hartford, CT.

Duran, R. L., & Zakahi, W. R. (1984). Competence or style: What's in a name. *Communication Research Reports, 1*, 42-47.

Hecht, M. L., Boster, F. J., & LaMer, S. (1989). The effect of extroversion and differentiation on listener-adapted communication. *Communication Reports, 2*(1), 1-8.

Reviewer: Judith M. Dallinger.

The Communication Behaviors Inventory

Authors/Developers: S. P. Morreale, D. D. Morley, & J. G. Naylor. S. P. Morreale, Center for Excellence in Oral Communication, Univ of Colorado at Colorado Springs, Austin Bluffs Pky, Colorado Springs CO 80933-7150.

Source: Morreale, S. P, Morley, D. D., & Naylor, J. G. (1991, November). *The CBI: Development of a grounded-theory communication competency measurement tool for large scale assessment*. Paper presented at the meeting of the Speech Communication Association, Atlanta.

Cost: Instrument not yet available; still under development.

Intended context or behavior: Overall communication competence.

Intended academic level: College.

Time required: 30-40 minutes.

Text description: Subjects respond to 170 semantic differential (1-7 scale) probes of their behaviors in 26 different communication situations. Subjects self-report their behaviors in public speaking, interpersonal, intrapersonal, and organizational contexts.

Culture and gender bias: Information not available.

Validity

Face/content: Focus groups, personal interviews, and telephone interviews were used to identify communication situations relevant to college subjects. Nine communication educators then evaluated and refined the items.

Evaluative Reactions

Practicality and usefulness: When completed, the instrument should provide a very thorough report of the individual's perception of his/her communication abilities.

Reviewer: Ellen A. Hay.

Communication Competency Self-Report Questionnaire

Author/Developer: R. B. Rubin. School of Communication Studies, Kent State Univ, P. O. Box 5190, Kent OH 44242-0001.

Source: Rubin, R. B. (1985). The validity of the communication competency assessment instrument. *Communication Monographs, 52,* 173-185.

Cost: Information not available.

Intended context or behavior: Overall self-perception of communication competence.

Intended academic level: College.

Time required: 15-20 minutes for administration and scoring.

Text description: The CCSR includes 19 statements related to an individual's perception of his/her public speaking, interaction, and listening abilities. Respondents indicate on a 5-point Likert-type scale the degree to which each statement describes their abilities. Half of the items are reversed.

Scoring description: The negatively worded items are reversed, and the 19 items are totaled. The article does not offer an interpretation of the scores. Means for each of the 19 items are given.

Norm/criterion reference populations: College students (N = 41) were used in developing this instrument.

Culture and gender bias: Information not available.

Validity

Concurrent: The CCSR was appropriate when correlated with the Personal Report of Communication Apprehension.

Face/content: The CCSR was developed as a parallel to the items on the Communication Competency Assessment Instrument which reflects the 19 functional communication competencies approved by the SCA.

Construct: Only two of the items on the CCSR were significantly correlated with the CCAI. Subjects tended to overestimate their own abilities. Instructors ratings of the respondents tended to be closer to the respondents' self-ratings than did the ratings of unknown observers.

Reliability

Internal consistency: An alpha coefficient of .87 was reported.

Evaluative Reactions

Practicality and usefulness: The 19 items provide information about the subject's perception of communication abilities in several situations. As a pre/post measure for a course or major, it may help to identify specific changes in perceptions.

Clarity of instructions: The directions are very clear.

Standardization in administration and evaluation: Administration and scoring are highly standardized.

Primary strengths: The CCSR helps subjects consider how their communication may change in particular situations. It is very quick and easy to administer.

Primary weaknesses: As validity studies indicate, self-perception may not be an accurate measure of communication competence. This particular instrument has not been very thoroughly tested. It was primarily designed to study the CCAI.

Overall adequacy: While it would be difficult to generalize too much from this instrument, it would serve as a good beginning for considering communication competence. When coupled with other measures, it might be very effective.

Additional References

Rubin, R. B. (1983). *Conceptualizing communication competence: Directions for research and instruction.* Kent, OH: Kent State University. (ERIC Document Reproduction Service No. 248 545)

Rubin, R. B. (1983, November). *Conceptualizing communication competence: Directions for research and instruction.* Paper presented at the meeting of the Speech Communication Association, Washington, DC.

[For more information, see the "Additional References" section of the *Communication Competency Assessment Instrument.*]

Reviewer: Ellen A. Hay.

Communication Flexibility Measure

Authors/Developers: M. M. Martin, & R. B. Rubin. M. M. Martin, Communication Dept, Mid-Michigan Comm College, Harrison MI 48625; R. B. Rubin, School of Communication Studies, Kent State Univ, P. O. Box 5190, Kent OH 44242-0001.

Source: Martin, M. M., & Rubin, R. B. (1994). Development of a communication flexibility measure. *Southern Communication Journal, 59*, 171-178.

Cost: Free/public domain.

Intended context or behavior: Flexibility in adapting communication behaviors to differing situations.

Intended academic level: College/Adult.

Time required: 15-20 minutes.

Text description: Students respond to 14 scenarios by indicating on a 5-point scale how closely the behaviors described in the scenario resemble their own. The scenarios involve acquaintances and friends/family in formal and informal interpersonal, group, and public settings.

Scoring description: Scoring on several items is reversed, and a total is computed. A mean of 49.99 (sd = 6.87), a median of 50, and a range between 21-65 were reported.

Norm/criterion reference populations: College students.

Culture and gender bias: Information not available.

Validity

Face/content: Scenarios describe a number of different situations and circumstances.

Construct: The instrument was developed separately from but related to communication adaptability, rhetorical sensitivity, and social desirability. Correlations on the Social Desirability Scale and Communication Adaptability Scale confirmed this premise. Results on the Rhetsen did not confirm a relationship with rhetorical sensitivity.

Reliability

Internal consistency: An alpha coefficient of .70 and a split-half correlation of .71 were reported.

Evaluative Reactions

Practicality and usefulness: This instrument would be useful in introducing and demonstrating the flexibility construct.

Clarity of instructions: The instructions are very clear.

Standardization in administration and evaluation: Instructions apply to all respondents.

Primary strengths: The scenarios in this measure are interesting and varied. They will serve as good discussion starters.

Primary weaknesses: The usefulness of this instrument will be further enhanced when a relationship is established with communication competence.

Overall adequacy: This instrument is very effective in encouraging students to think about the contextual nature of communication.

Additional References

Parks, M. R. (1985). Interpersonal communication and the quest for personal competence. In M. L. Knapp & G. R. Miller (Eds.), *Handbook of interpersonal communication* (pp. 171-201). Beverly Hills: Sage.

Wiemann, J. M. (1977). Explication and test of a model of communication competence. *Human Communication Research, 3*, 195-213.

Reviewer: Ellen A. Hay.

Focal Person's Communications Survey Colleague Questionnaire

Authors/Developers: R. Klauss, & B. M. Bass.

Address: Information not available.

Source: Klauss, R., & Bass, B. M. (1982). *Interpersonal communication in organizations*. New York: Academic Press.

Cost: Information not available.

Intended context or behavior: Communicator Style.

Intended academic level: College/Adult.

Time required: 30-45 minutes.

Text description: A 25-item, 7-point, Likert-type scale that includes a self-report and a report of colleagues, which are used together. Dimensions include five communication styles, three aspects of credibility, and three outcomes.

Scoring: Information not available.

Norm/criterion reference populations: Means summarized in the text are derived from the Navy, a high-technology firm, a traditional production facility, a military base, and a social service agency.

Culture and gender bias: Information not available.

Validity

Construct: Scores for descriptions of the focal person and self-report scales converge.

Reliability

Test/retest: Ranged from .56 to .92 for different dimensions.

Internal consistency: Above .80.

Evaluative Reactions

Primary weaknesses: Insufficient information has been released to the public regarding this instrument.

Additional References

Bass, B. M., & Avolio, B. J. (1989). Potential biases in leadership measures: How prototypes, leniency and general satisfaction relate to ratings and rankings of transformational and transactional leadership constructs. *Educational and Psychological Measurement, 49*, 509-527.

Downs, C. W., Archer, J., McGrath, J., & Stafford, J. (1988). An analysis of communication style instrumentation. *Management Communication Quarterly, 1*, 543-571.

Reviewer: Judith M. Dallinger

Assessment of Communication Competency & English Speaking Skills (I ACCESS)

Author/Developer: N. L. Flores, Golden West College, 15744 Golden West St, Box 2748, Huntington Beach CA 92647. 714-892-7711.

Source: Information not available.

Cost: $25 for I ACCESS test and User's Manual with sample assessment forms and videotaped anchors.

Intended context or behavior: Speaking, listening, pronunciation, interpersonal, and group.

Intended academic level: College/Adult.

Time required: 3 hours training/ administration, 3-5 minutes per student, 5-6 minutes to view responses, 1 hour to view 20 tapes.

Text description: Examinees (a) see a visual prompt depicting culturally diverse speakers and listeners interacting in culturally diverse job situations; (b) read the question while listening to a three-part, job-related question; and (c) answer the question as it is recorded. After 20 responses, cassette, appointment sheet, and assessment forms are given to assessor.

Scoring description: The assessment form allows the assessor to outline responses and perceived mispronunciations. The form is also divided into seven competencies: purpose/focus, organization, examples, vocabulary/grammar, pronunciation/emphasis, nonverbal congruency, and use of time allotted. Each competency is attended by five criteria (scored 1-5) that can be observed as behaviors demonstrated. Seven competencies times five criteria yield 35 total points.

Norm/criterion reference populations: The first test involved 118 freshmen and sophomores at Golden West College.

Culture and gender bias: Information not available. [Instrument is still in the testing process.]

Validity:

Information not available. [Instrument is still in the testing process.]

Reliability

Information not available. [Instrument is still in the testing process.]

Evaluative Reactions

Practicality and usefulness: This instrument can be useful as a cultural competency instrument/test.

Clarity of instructions: It is easy to understand and follow.

Primary strengths: The focus is primarily on testing and developing speaking and listening competencies in speakers of English as a second language.

Primary weaknesses: The instrument needs to be validated for reliability.

Overall adequacy: This tool has strong potential.

Additional References

Flores, N. L. (1992). *Multicultural communication: Assessing, facilitating, and evaluating competency-based oral interactions*. Huntington Beach, CA: Golden West College. (ERIC Document Reproduction Service No. ED 351 734)

Flores, N. L. (1992, November). *Multicultural communication: Assessing, facilitating, and evaluating competency-based oral interactions*. Short course presented at the meeting of the Speech Communication Association, Chicago.

Gomez, A. M., Ricillo, S., Flores, N. L., Cooper, P., & Starosta, W. (1994). Intercultural communication competence: Developing assessment instruments. In S. P. Morreale, M. Brooks, R. Berko, & C. Cooke (Eds.), *1994 SCA summer conference proceedings and prepared remarks* (pp. 389-396). Annandale, VA: Speech Communication Association.

Reviewer: Adelina M. Gomez.

Norton Communicator Style Measure (CSM)

Author/Developer: R. W. Norton.

Address: Information not available.

Source: Norton, R. W. (1978). Foundation of a communicator style construct. *Human Communication Research, 4*, 99-111.

Cost: Information not available.

Intended context or behavior: Communicator style.

Intended academic level: College/Adult.

Time required: 8-10 minutes.

Text description: A general self-report instrument that contains nine subscales that are conceptualized as independent variables, and one subconstruct that is characterized as a dependent variable. There are 51 items: 5 each for the 10 subconstructs, using a 4-point Likert-type response. A sixth item relating to the individual's communicator image for each subconstruct asks each subject to rank order one's self-perception in relation to the typical communicator.

Scoring description: Scores for each subconstruct are summed from appropriate items, as described.

Norm/criterion reference populations: Research is based on a sample of 1,086 college students.

Culture and gender bias: Information not available.

Validity

Face/content: Items on the instrument appear to be appropriate for the construct.

Reliability

Internal consistency: Alpha reliabilities are from .37 for friendly, .56 for animated, .57 for attentive, .65 for contentious, .68 for dramatic, .69 for impression leaving, .69 for open, .71 for related, .72 for communicator image, and .82 for dominant.

Evaluative Reactions

Practicality and usefulness: The scale is easy to administer.

Clarity of instructions: The directions are clear.

Primary strengths: The instrument has been used in many research studies.

Primary weaknesses: Questions remain about the structure of the scale.

Overall adequacy: This has some potential to be developed as an assessment technique.

Additional References

DeWine, S., & Pearson, J. C. (1985). *The most frequently used self-report instruments in communication*. Athens: Ohio University. (ERIC Document Reproduction Service No. ED 260 479).

DeWine, S., & Pearson, J. C. (1985, May). *The most frequently used self-report instruments in communication*. Paper presented at the meeting of the International Communication Association, Honolulu.

Downs, C. W., Archer, J., McGrath, J., & Stafford, J. (1988). An analysis of communication style instrumentation. *Management Communication Quarterly, 1*, 543-571.

Forsyth, P. B., & Boshart, D. (1985). *Leadership style and principal communication: A preliminary investigation*. New Brunswick, NJ: Rutgers University. (Eric Document Reproduction Service No. 263 664)

Forsyth, P. B., & Boshart, D. (1985, March). *Leadership style and principal communication: A preliminary investigation*. Paper presented at the meeting of the American Educational Research Association, Chicago.

Norton, R. W. (1983). *Communicator style: Theory, applications and measures*. Beverly Hills, CA: Sage.

Norton, R. W. (1986). Communicator style in teaching: Giving good form to content. *New Directions for Teaching and Learning, 26*, 33-40.

Notarantonio, E. M., & Cohen, J. L. (1990). The effects of open and dominant communication styles on perceptions of the sales interaction. *Journal of Business Communication, 27*(2), 171-184.

Slaughter, G. Z., Jorgensen, J. D., & Petelle, J. L. (1991). *The congruity of communicator style expectations and experiences and its influence upon subordinates' satisfaction with supervision*. Belvue, NE: Belvue College. (ERIC Document Reproduction Service No. 340 057)

Slaughter, G. Z., Jorgensen, J. D., & Petelle, J. L. (1991, November). *The congruity of communicator style expectations and experiences and its influence upon subordinates' satisfaction with supervision*. Paper presented at the meeting of the Speech Communication Association, Atlanta.

Reviewer: Judith M. Dallinger.

Openness Scale

Authors/Developers: R. W. Norton, & B. M. Montgomery. B. M. Montgomery, Academic Affairs Office, Univ of New Hampshire, Durham NH 03824.

Source: Norton, R. W., & Montgomery, B. M. (1982). Style, content, and target components of openness. *Communication Research*, *9*, 399-431.

Cost: Free/public domain.

Intended context or behavior: Willingness to share information about self with others.

Intended academic level: College.

Time required: 15-20 minutes.

Text description: This scale is structured around three dimensions of openness: style, content, and target. The first section contains 19 items related to style. Students respond on a 5-point Likert-type scale how closely their behaviors resemble those described. The second section includes 17 general topics. Students indicate their degree of openness to discuss each topic on a 7-point scale. The third section specifies six different message recipients. Students indicate on a 9-point scale the degree of openness they are likely to show each target.

Scoring description: Scoring instructions are not included in the article.

Norm/criterion reference populations: College students.

Culture and gender bias: Information not available.

Validity

Face/content: The items in Part I of the scale were drawn from research in the field. The other two parts included items that reflected the range of topics and likely targets.

Predictive: Individual items in the three sections were correlated appropriately with self-reports of communication ability, openness, willingness to disclose, and length of time it takes to develop familiarity.

Construct: The construct of "openness" is defined in relation to research on style, content, and target. Factor analyses were used to identify items that define the construct.

Reliability

Information not available.

Evaluative Reactions

Practicality and usefulness: The instrument helps introduce some of the dimensions of self-disclosure.

Clarity of instructions: The directions for each section seem clear.

Standardization in administration and evaluation: Administration procedures are fairly routine.

Range of responses or degree of competence: No interpretative information is provided in the article.

Primary strengths: In addition to focusing on self-disclosure in general, the instrument also encourages students to consider some of the aspects of communication style.

Primary weaknesses: Since each subscale has a different scoring protocol, the measure is somewhat cumbersome. This measure needs to be correlated with other measures of self-disclosure.

Overall adequacy: It accomplishes its intended objective.

Additional References

Montgomery, B. M. (1980). Trait interactionist and behavior assessment variables in open communication. *Communication Research-An International Quarterly*, *7*, 479-494.

Montgomery, B. M. (1981). The form and function of quality communication in marriage. *Family Relations*, *30*, 21-29.

Reviewer: Ellen A. Hay

Self-Perceived Communication Competence Scale

Authors/Developers: J. C. McCroskey, & L. L. McCroskey. J. C. McCroskey, Communication Studies Dept, West Virginia Univ, Morgantown WV 26506.

Source: McCroskey, J. C., & McCroskey, L. L. (1988). Self-report as an approach to measuring communication competence. *Communication Education*, 5, 108-113.

Cost: Information not available.

Intended context or behavior: Overall perception of competence as a communicator.

Intended academic level: College/Adult.

Time required: 10-15 minutes to complete and score.

Text description: The 12 items ask respondents to rate their communication competence on a scale of 0-100 in four contexts (public speaking, large meeting, small group, interpersonal) and with three types of receivers (stranger, acquaintance, friend).

Scoring description: Respondents add designated items for a score on each context and type of receiver. A global score is computed for overall competence and subscores on context and receiver. The article reports mean global and subscale scores that could be used for comparison.

Norm/criterion reference populations: Undergraduate students (N = 344) were used in developing this instrument.

Culture and gender bias: Information not available.

Validity

Face/content: Items were based upon the McCroskey definition of communication competence, which is adequate ability to pass along or give information; the ability to make known by talking or writing.

Construct: The article discusses the role of self-report instruments as valid measures of respondent perceptions.

Reliability

Internal consistency: Reliability estimates for subscores ranged from .44 to .87. Global score reliability was estimated at .92.

Evaluative Reactions

Practicality and usefulness: It provides an indication of the respondent's perceptions of his/her communication competence. It does not measure actual competence. It could be used at the beginning and ending of a course or major to demonstrate changes in one's perception of communication abilities.

Clarity of instructions: Instructions for completing and scoring the scale are very straightforward.

Standardization in administration and evaluation: It should be relatively easy to administer and score this measure consistently.

Range of responses or degrees of competence: Labels are not attached to scores. While a mean score is given and could be used as a basis for comparison, no judgments are made on the perception of competence.

Primary strengths: This measure is very efficient. It provides a quick, easy indication of growth.

Primary weaknesses: Not much testing has been done with regard to the reliability, validity, and bias of this measure.

Overall adequacy: It provides another way of measuring the development of communication abilities. It would be particularly good as a "discussion starter."

Additional References

McCroskey, J. C., & McCroskey, L. L. (1983). *Don't speak to me in English: Communication apprehension in Puerto Rico*. Morgantown, WV: West Virginia University. (ERIC Document Reproduction Service No. 268 592)

McCroskey, J. C., & McCroskey, L. L. (1983, December). *Don't speak to me in English: Communication apprehension in Puerto Rico*. Paper presented at the Speech Communication Association of Puerto Rico convention, San Juan.

McCroskey, J. C., & McCroskey, L. L. (1985). Don't speak to me in English: Communication apprehension in Puerto Rico. *Communication Quarterly*, 33(3), 185-192.

McCroskey, J. C., & McCroskey, L. L. (1986). *Self-Report as an approach to measuring communication competence*. Morgantown, WV: West Virginia University. (ERIC Document Reproduction Service No. 268 592)

McCroskey, J. C., & McCroskey, L. L. (1986, April). *Self-Report as an approach to measuring communication competence*. Paper presented at the meeting of the Central States Speech Association, Cincinnati.

Reviewer: Ellen A. Hay.

Source Credibility Scale

Author/Developer: J. C. McCroskey. Communication Studies Dept, West Virginia Univ, Morgantown WV 26506-3905.

Source: Rubin, R. B. (1994). Source credibility scale-McCroskey. In R. B. Rubin, P. Palmgreen, & H. E. Sypher (Eds.), *Communication research measures: A sourcebook* (pp. 332-337). New York: Guilford Press.

Cost: Free/public domain.

Intended context or behavior: Perceptions of source credibility.

Intended academic level: College/Adult.

Time required: 5-10 minutes.

Text description: McCroskey has designed two semantic differentials that are intended to measure various dimensions of source credibility. Subjects are asked to identify an individual to evaluate. One measure focuses on authoritativeness and character, each with six items. Another scale measures sociability, extroversion, competence, composure, and character, with three items each. Both scales use 7-point semantic differentials.

Scoring description: Some items are reversed and responses are totaled. To study various dimensions, subscales can be compared. Interpretative ranges were not given.

Norm/criterion reference populations: College students.

Culture and gender bias: Information not available.

Validity

Predictive: Scores on similar versions of the measure adapted for teachers predicted whether students would be likely to enroll in a subsequent class with a particular teacher.

Construct: The scale was found to correlate appropriately with various qualities typically associated with credibility, including speech rate, humor, status, evidence, appearance, and powerful language.

Reliability

Alternate forms: : The Likert-type versions of the scales do not seem to be as reliable as those using the semantic differential.

Internal consistency: Both scales report sufficient levels of consistency within each dimension.

Evaluative Reactions

Practicality and usefulness: These scales are very quick and easy to use.

Clarity of instructions: Directions are easy to follow.

Standardization in administration and evaluation: Both scales ask students to focus upon a particular individual.

Primary strengths: This is a frequently used measure and would be effective in demonstrating the various dimensions of credibility. Students might also consider changes in their own credibility.

Primary weaknesses: It would be interesting to have some comparative/interpretative data.

Overall adequacy: This instrument is well-suited for instructional purposes.

Additional References

McCroskey, J. C. (1966). Scales for the measurement of ethos. *Speech Monographs, 33*, 65-72.

McCroskey, J. C., Hamilton, P. R., & Weiner, A. M. (1974). The effect of interaction behavior on source credibility, homophily, and interpersonal attraction. *Human Communication Research, 1*, 42-52.

McCroskey, J. C., Holdridge, W., & Toomb, J. K. (1974). An instrument for measuring the source credibility of basic speech communication instructors. *Speech Teacher, 23*, 26-33.

McCroskey, J. C., & Young, T. J. (1981). Ethos and credibility: The construct and its measurement after three decades. *Central States Speech Journal, 32*, 24-34.

Reviewer: Ellen A. Hay.